RICE UNIVERSITY

An Alternative Approach to Differential Semblance Velocity Analysis via Normal Moveout Correction

by

Chao Wang

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Master of Arts

Approved, Thesis Committee:

Dr. William W. Symes, Chair
Noah Harding Professor of Computational and
Applied Mathematics

Dr. Matthias Heinkenschloss
Professor of Computational and Applied Mathematics

Dr. Yin Zhang
Professor of Computational and Applied Mathematics

Houston, Texas

March, 2010

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

An Alternative Approach to Differential Semblance Velocity Analysis via Normal Moveout Correction

by

Chao Wang

This thesis develops a new computation of the objective function and gradient for normal moveout-based differential semblance (DS). The DS principle underlies a class of algorithms for seismic velocity analysis. The simplest variant of DS is based on a drastic approximation to the scattering of waves, called "normal moveout" (NMO) in the seismic literature. This simple NMO-driven DS algorithm is very fast relative to other variants based on more faithful approximations to wave physics, but nonetheless accurate enough to be used to process field data. A recent implementation of NMO-based DS demonstrated these capabilities, but it also exhibited numerical irregularity which may have affected the stability of its velocity estimates. My alternative approach avoids interpolation noise that existed in previous work and so results in more stable numerical optimization.

Acknowledgements

I would like to express my deep and sincere gratitude to my academic advisor, Professor William W. Symes. He has been helping me accumulate research skills, as well as build up confidence in conducting research. His profound knowledge and upbeat personality have greatly influenced me. Without his guidance, I might not have found research so enjoyable. I owe many thanks to him for his encouragement, patience, and guidance. I really take pride in working with him.

I would also like to extend my deepest appreciation to Ivan Vasconcelos for his leadership and mentorship during the summer of 2008 at ION/GXT. He helped me extend my knowledge base regarding my thesis and his guidance will always be remembered. I also thank Omer Alpak for his mentorship during the summer of 2009 at Shell. He helped me expand my knowledge in the geoscience industry.

I also thank Dr. Hewitt for her support and help with revising my thesis on several occasions. I owe many thanks to her for her suggestions of my usage of the English Language and helpful recommendations for my presentations. I learned many communication skills from her which are very useful not only for my presentations but also for my whole life.

I am grateful for Professor Heinkenschloss, who was my mentor during my first year at CAAM and gave me a lot of support and guidance. I also would like to thank Professor Zhang and Professor Yin, who also provided me plenty of help.

I am grateful for all the professors who have provided me knowledge and expertise during my undergraduate and graduate studies.

Finally, I would like to thank my parents and all my dear friends.

Contents

List of Figures

Chapter 1

Introduction

Seismic data typically consists of recordings of waves that are generated by a controlled source at the surface and propagate through the subsurface of the earth. Reflected waves are recorded by seismic recording equipment at the surface and converted to digital signals. Various parameters describing the rock properties are sought from these recorded data. One particularly critical parameter is the velocity at which the waves propagate through a medium. Velocity analysis is the process of estimating wave velocity from the recorded seismic data. Successful seismic data analysis requires accurate knowledge of the velocity.

Velocity analysis can be considered as an inverse problem. The aim is to automatically reconstruct the velocity model of the earth from seismic data recorded at the surface. The most straightforward strategy is data-fitting inversion. It can be considered as an optimization problem in the form of a misfit function to be minimized.

It involves matching of recorded and simulated seismic data using an approximated velocity. However, this problem is highly nonlinear and has many spurious local minima. Another strategy is migration velocity analysis. Migration is a generic name for velocity-dependent process which maps recorded seismic data to a seismic image, given a suitable velocity. A seismic image is a representation that depicts the detail of subsurface structure. Migration velocity analysis exploits the redundancy in the data. If the redundant data is sorted into "smaller" data sets, each of these data sets can form a subsurface image independent of the other subsets, even though all the subsets represent the same part of the earth. Thus seismic data contains enough information to produce multiple independent images. Migration velocity analysis involves measuring and correcting inconsistencies between redundant images that indicate velocity inaccuracies. Differential semblance, which is an operator used to evaluate the image consistency, is required for differential semblance migration velocity analysis, which have been studied extensively [Symes, 2008].

Among many variants of differential semblance, differential semblance via normal moveout correction is the simplest form. Normal moveout (NMO) is defined as the difference in reflected wave traveltimes from a horizontal reflecting surface due to variations in the source-receiver distance. NMO correction is a function of traveltime (the elapsed time for wave to travel from its source to a given reflector and return to a receiver) and half offset (half of the distance from source to receiver) that can be used to compensate the delay in traveltime when sources and receivers are offset from each

other. The seismic data is gridded in the traveltime - half offset domain and the image is in the domain of traveltime at zero-offset and half offset. NMO correction maps the seismic data to an image by velocity-dependent change of variables. Therefore it is inevitable to involve numerical interpolation.

A previous numerical implementation of objective function is based on this change of variables, so I name this previous DS approach DS-CV [Li and Symes, 2007]. Large interpolation error of the oscillatory seismic data, which may lead to the wrong velocity estimation, is inevitable in the numerical implementation if the time sampling intervals are too far apart. The DS-CV objective function exhibits spurious local minima in my numerical examples. In order to avoid the error caused by the interpolation of the oscillatory seismic data, I proposed a new discretization method to approximate the same continuous DS objective function as DS-CV. This new approach of the implementation is based on the chain rule, so I name my DS approach DS-CR.

My claim is that DS-CR only involves interpolation of the smooth function by applying the chain rule to compute the objective function. The DS-CR objective function does not exhibit local minima, which enables DS-CR to be an effective and accurate approach to estimate the velocity if the seismic data has small offset sampling rate or contains low frequency data. My numerical results show that, unlike DS-CV, the DS-CR objective function always has one global minimum (no local minimum). I also test for different offset intervals and frequencies, and conclude that DS-CR should

be effective if the sufficient conditions, which I computed in chapter 3, are satisfied. From the numerical experiments of synthetic and real data, I can conclude that DS-CR will be effective if and only if the seismic data has small offset sampling rate or contains low frequency data. If the offset sampling rate is too big or the data with low frequency are missing, DS-CR will have large discretization error to approximate partial derivative with respect to half offset and lead to premature termination during optimization. A better algorithm should be employed to compute the offset derivative. I propose an alternative option in numerical implementation which combines DS-CV and DS-CR in the end.

In chapter 2, I provide the theory and some background information for differential semblance. In chapter 3, I compare the advantages and disadvantages of different approaches to differential semblance. I also derive the numerical methods used to implement DS-CR and estimate the error bound for the numerical computation of the objective function. The implementation is done within the framework of the Rice Vector Library (RVL) and the Seismic Unix (SU) packages. RVL is a mathematical library which provides simple access to some very powerful optimization algorithms. I employ the Limited Memory Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm (LBFGS) using the RVL framework to solve the optimization problem. SU is intended as an extension of the Unix operating system. This package is not restricted to seismic processing tasks. A broad suite of wave-related processing can be done with SU. In my thesis, SU provides me with a powerful tool to generate, analyze and

process seismic data.

Chapter 4 and chapter 5 present the applications of my technique to the synthetic and real data. I outline the results from the experiments performed. The data sets on which I perform DS-CR are 2D synthetic data generated by Seismic Unix and 2D marine data from the North Sea. I also compare my results with the results produced by DS-CV in order to demonstrate pros and cons of both approaches.

In chapter 6, I discuss of the difficulties encountered in the numerical tests and the suggestions for future development. This thesis ends with conclusions of my achievement and discussions of the pros and cons of DS-CR compared to DS-CV.

Chapter 2

Theory

Differential semblance (DS) via normal moveout correction (NMO) is the simplest form among all differential semblance approaches. This simplest variant of differential semblance is based on drastic assumptions and approximations to the scattering of waves. Model assumptions include that the medium, through which the wave propagates, is an acoustic layered medium with constant density, and that the energy source has compact support at a single point. The theoretical foundations of DS via NMO are three approximations: partial linearization, high frequency asymptotics, and hyperbolic moveout approximations, which will be explained in detail in this chapter.

2.1 Geophysical Terminology

This thesis uses some geophysical terms because this subject is also a topic in exploration geophysics. I provide these definitions for all the geophysical terms used in this thesis for people who are not familiar with geophysics. Some of the definitions below come from the Schlumberger online oilfield glossary. It is easy to understand this terminology by referring to Figure 2.1, together with the definitions.

CMP: Common midpoint which is the point on the surface halfway between the source and the receiver that may be shared by numerous source-receiver pairs. Common midpoint is shown as the point M in Figure 2.1.

Offset: The horizontal distance from source to receiver in the surface seismic acquisition. Half offset is denoted by h. In Figure 2.1, offset is the distance between source S_2 and receiver (geophone) R_1 and half offset is the distance between common midpoint M and receiver (geophone) R_1.

Trace: A recording of the earth's response to seismic energy passing from the source, through subsurface layers, and back to the receiver.

Ray: Geometric optics describes wave propagation in terms of rays. The ray can be used to predict the wave path and is perpendicular to the wavefront. The ray bends when it reaches the reflector, which is the interface between two different media. It is shown in Figure 2.1 that the ray (wave path) starts from source

S_2, reflects at depth point O on the interface between medium 1 and medium 2, and is received by receiver R_1.

Event: An appearance of seismic data as a reflection, diffraction, refraction or other similar feature produced by an arrival of seismic energy. This thesis focuses on the reflection event. An event can be a single wiggle within a trace, or a consistent lining up of several wiggles over several traces. An event in a seismic section can represent a geological interface, such as a fault, or change in lithology.

Primary reflection: Seismic events whose energy has been reflected once. In Figure 2.1, primary reflection occurs when the ray reaches the reflector, which is the interface between medium 1 and medium 2. It satisfies the reflection law, which says that the incident angle $\theta 1$ equals to the reflected angle $\theta 2$. Multiples, in contrast, are events whose energy has been reflected more than once.

Traveltime: The duration of the passage of a primary reflection signal from the source through the earth and back to the receivers. Also called two-way traveltime which is denoted by t. In Figure 2.1, t is the time when waves travel along the ray path from source S_2 to depth point O then back to receiver R_1. Zero-offset two-way traveltime is that of a normal-incidence wave with zero offset which is denoted by t_0.

Moveout: The difference in the traveltimes of a reflected wave measured by receivers

at two different offset locations.

NMO: Normal moveout which is the moveout caused by the separation between a source and a receiver in the case of a flat reflector. NMO can be considered as two-way traveltime delay relative to the zero-offset two-way traveltime.

Gather: Display of seismic traces that share an acquisition parameter, such as common midpoint gather, which contains traces having a common midpoint.

Image: A representation that depicts the detail of subsurface structure of the earth.

Migration: A velocity dependent imaging process which maps recorded seismic data to a seismic image.

Isotropic medium: A medium in which the waves travel at equal speed independently of which direction they travel.

2.2 Acoustic Model of Seismic Reflection

Seismic surveys are performed to gather the information about the geology of an oil or gas field. Figure 2.1 shows the geometry of a simple seismic survey. A controlled seismic source of energy $f(\mathbf{x}, t; \mathbf{x_s})$, such as dynamite on land or an airgun in water, generates the waves. The waves propagate away from the source and travel through the subsurface medium of the earth. The reflected waves are recorded using seismometers, such as geophones on land or hydrophones in water. Those instruments

convert the pressure changes into electrical signals.

Body waves are the waves that travel through the interior of the Earth. There are two types of body waves: P-waves and S-waves. P-waves (primary waves) are longitudinal waves, which means that the particles in the medium have vibrations along or parallel to the direction of wave propagation. S-waves (shear waves) are transverse waves, which means that the particles vibrate perpendicularly to the direction of wave propagation. The acoustic wave equation governs the motion of P-waves through a material medium resulting from an energy source, while the elastic wave equation governs the motion of P-waves and S-waves through the medium. The acoustic wave equation, based on the acoustic assumption (S-wave velocity equals zero), is much simpler to solve than elastic wave equation. This assumption is physically reasonable since the water layer is free of S-waves in a marine survey, though it is not true for the layers of sedimentary rocks under the water layer. The acoustic wave equation yields good kinematic approximation to the wave propagation in the real physical models.

The acoustic wave equation is a second order partial differential equation. I assume that the density is constant (equal to one), and the source has point support for simplification. The assumption that density is invariant with depth simplifies the model representation, but is not always valid. The reason that I can neglect the density variation is that the density gradient usually has the same sign as the velocity gradient and density variations are usually much smaller than velocity variations, thus

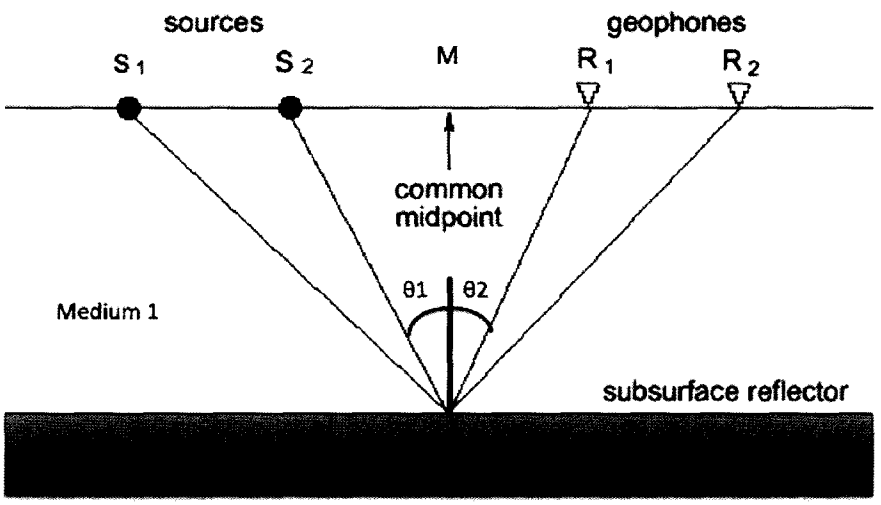

Figure 2.1: Geometry of a simple seismic survey. Two sources S_1,S_2 and two receivers (geophones) R_1,R_2. Common midpoint (CMP) M. Incident angle $\theta 1$ and $\theta 2$. Medium 1 and medium 2 with two different velocities. Subsurface reflector is the interface between medium 1 and medium 2. Depth point O is the reflection point of the ray paths.

it still provides an useful wave equation. The point source can be written as

$$f(\mathbf{x}, t; \mathbf{x_s}) = \omega(t)\delta(\mathbf{x} - \mathbf{x_s}),$$

where $t \in \mathrm{R}$ is time, $\mathbf{x} \in \mathrm{R}^3$ is the position vector, $\mathbf{x_s} \in \mathrm{R}^3$ is the position of the point source, and $\omega(t)$ is the source time function. This point source assumption is reasonable when the spatial extend of the source is much smaller than a typical data wavelength. Under these assumptions, the acoustic wave equation can be written as

$$\frac{1}{c^2(\mathbf{x})}\frac{\partial^2 p}{\partial t^2}(\mathbf{x}, t; \mathbf{x_s}) - \nabla^2 p(\mathbf{x}, t; \mathbf{x_s}) = \omega(t)\delta(\mathbf{x} - \mathbf{x_s}), \tag{2.1}$$

with initial and boundary conditions. Here $t \in \mathrm{R}$ is time, $\mathbf{x} \in \mathrm{R}^3$ is the position vector, $\mathbf{x_s} \in \mathrm{R}^3$ is the position of the point source, $\omega(t)$ is the source time function, $c(\mathbf{x})$ is the P-wave velocity, and $p(\mathbf{x}, t; \mathbf{x_s})$ is the acoustic pressure. I also assume that the medium is in its equilibrium state (of zero pressure) for negative time,

$$p(\mathbf{x}, t; \mathbf{x_s}) \equiv 0, t < 0.$$

Physical boundaries may imply boundary conditions. But I will ignore these since they do not alter the general principles presented below.

Forward map:

The Green's function $G(\mathbf{x}, t; \mathbf{x_s})$, which is observed at $\mathbf{x}$ subject to source located at $\mathbf{x_s}$, is given by changing $\omega(t)$ to $\delta(t)$. Then the Green's function $G(\mathbf{x}, t; \mathbf{x_s})$ satisfies

$$\left(\frac{1}{c^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) G(\mathbf{x}, t; \mathbf{x_s}) = \delta(t)\delta(\mathbf{x} - \mathbf{x_s}) \quad , \tag{2.2}$$

$$G(\mathbf{x}, t; \mathbf{x_s}) \equiv 0 \quad , \quad t < 0. \tag{2.3}$$

Convolve the above equation 2.2 with the source time function $\omega(t)$, it yields

$$\int \omega(t-\hat{t}) \left(\frac{1}{c^2(\mathbf{x})} \frac{\partial^2}{\partial \hat{t}^2} - \nabla^2 \right) G(\mathbf{x}, \hat{t}; \mathbf{x_s})\, d\hat{t} = \int \omega(t-\hat{t}) \delta(\hat{t}) \delta(\mathbf{x}-\mathbf{x_s})\, d\hat{t}.$$

Since

$$\frac{\partial}{\partial t}(\omega *_t G) = \frac{\partial \omega}{\partial t} *_t G = \omega *_t \frac{\partial G}{\partial t},$$

apply the property of the delta function and get

$$\left(\frac{1}{c^2(\mathbf{x})} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \int \omega(t-\hat{t}) G(\mathbf{x}, \hat{t}; \mathbf{x_s})\, d\hat{t} = \omega(t) \delta(\mathbf{x}-\mathbf{x_s}). \tag{2.4}$$

This means that the solution to the equation 2.1, which is the pressure field, is the source convolved with the Green's function

$$p(\mathbf{x}, t; \mathbf{x_s}) = \omega(t) *_t G(\mathbf{x}, t; \mathbf{x_s}).$$

It allows us to define the forward mapping operator S. Operator S depends on velocity $c(\mathbf{x})$ and can be computed from the acoustic wave equation 2.1. It maps velocity c into pressure field at a number of receiver locations $\mathbf{x_r}$ and times t. The pressure field can be regarded as the predicted data

$$S[c] = p|_{Y=\{(\mathbf{x_r}, t; \mathbf{x_s})\}}.$$

Seismic inverse problem:

Given observed seismic data S^{obs}, find velocity c so that

$$S[c] \simeq S^{obs}.$$

Even with all the modeling assumptions I made for simplification, this seismic inverse problem is still very difficult to solve due to three major obstacles. The first

one is computational intensity at industrial or scientific scale. Wavefields modeling and other related computational tasks are required by inversion. The computational scale is as large as up to Tbytes, Pflops. But this computational obstacle is gradually resolved by improvements in computer performance and algorithm efficiency. The convolutional model is one way to reduce the computation cost (to be explained in the next section).

The second obstacle is that the seismic inverse problem is usually ill-posed. A mathematical problem is called well-posed if it satisfies three conditions: 1) a solution exists; 2) the solution is unique; 3) the solution depends continuously on the data (well-conditioned). If one of these properties is violated, the problem is called ill-posed. A seismic inverse problem usually has non-unique solutions and/or is ill-conditioned. In practical applications, algorithms for well-posed problem may fail if the problem is ill-conditioned, since noisy data as well as round-off errors may cause large variations in the solution. In order to solve an ill-posed problem one has to use regularization methods, which in general terms replace the ill-posed problem by a family of nearby well-posed problems.

The other obstacle is nonlinearity. Since the relation $c \rightarrow p$ is nonlinear, the seismic inverse problem $p \rightarrow c$ is nonlinear as well. Nonlinear problems are intrinsically more difficult to solve than linear problems. Due to the complexity of the problem, complicated computational schemes are required and related computational cost is rapidly increased by applying these approaches to solve the nonlinear problem.

2.3 Partial Linearization

Even with all the simplifying assumptions stated in section 2.2, the recovery of $c(\mathbf{x})$ from seismic data is still extremely difficult mainly due to the nonlinear relation between the velocity coefficients c and the solution p of the acoustic wave equation. My first step is the linearization of the c and p relation. We can decompose c as a background velocity v and its first order velocity perturbation δv. Write

$$c(\mathbf{x}) = v(\mathbf{x}) + \delta v(\mathbf{x}), \tag{2.5}$$

and define the relative velocity perturbation (reflectivity) $r(\mathbf{x}) = \frac{\delta v(\mathbf{x})}{v(\mathbf{x})}$. Substitute equation 2.5 into equation 2.1 and replace p by $p + \delta p$. It yields

$$\left(\frac{1}{(v+\delta v)^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)(p+\delta p)(\mathbf{x}, t; \mathbf{x_s}) = \omega(t)\delta(\mathbf{x}-\mathbf{x_s}), \tag{2.6}$$

where p satisfies

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)p(\mathbf{x}, t; \mathbf{x_s}) = \omega(t)\delta(\mathbf{x}-\mathbf{x_s}). \tag{2.7}$$

Multiply both sides of equation 2.6 by $(v+\delta v)^2(\mathbf{x})$

$$\left(\frac{\partial^2 p}{\partial t^2} + \frac{\partial^2 \delta p}{\partial t^2} - (v+\delta v)^2(\mathbf{x})\nabla^2 p - (v+\delta v)^2(\mathbf{x})\nabla^2\delta p\right)(\mathbf{x}, t; \mathbf{x_s}) = (v+\delta v)^2(\mathbf{x})\omega(t)\delta(\mathbf{x}-\mathbf{x_s}). \tag{2.8}$$

Multiply both sides of equation 2.7 by $v^2(\mathbf{x})$

$$\left(\frac{\partial^2 p}{\partial t^2} - v^2(\mathbf{x})\nabla^2 p\right)(\mathbf{x}, t; \mathbf{x_s}) = v^2(\mathbf{x})\omega(t)\delta(\mathbf{x}-\mathbf{x_s}). \tag{2.9}$$

Subtracting equation 2.9 from equation 2.8, and keeping the first order terms (neglect the terms having δv^2 or $\delta v \nabla^2 \delta p$) yield

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\delta p(\mathbf{x}, t; \mathbf{x_s}) = \frac{2\delta v(\mathbf{x})}{v^3(\mathbf{x})}\frac{\partial^2 p}{\partial t^2}(\mathbf{x}, t; \mathbf{x_s}). \tag{2.10}$$

The linear approximation leads to

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\delta p(\mathbf{x}, t; \mathbf{x_s}) = \frac{2r(\mathbf{x})}{v^2(\mathbf{x})}\frac{\partial^2 p}{\partial t^2}(\mathbf{x}, t; \mathbf{x_s}). \tag{2.11}$$

My partial linearization is valid so long as the background velocity v is slowly-varying (smooth) relative to a typical data wavelength and its relative velocity perturbation r is oscillatory (rough) [Symes, 1998a]. There is no strict mathematical criteria for these arguments so far. Based on this linearized acoustic wave equation, I can define the partially linearized forward map $\hat{F}$ as below.

The partially linearized forward map:

$$\hat{F}[v]r = \delta p|_{Y=\{(\mathbf{x_r}, t; \mathbf{x_s})\}}.$$

Partial linearization means that the operator $\hat{F}$ depends nonlinearly on the smooth background velocity v and linearly on the oscillatory relative velocity perturbation r.

The partially linearized inverse problem (velocity analysis problem):

Given observed seismic data S^{obs}, find a smooth background velocity v and its oscillatory relative velocity perturbation r so that

$$S[v] + \hat{F}[v]r \simeq S^{obs}.$$

2.4 Convolutional Model

The partially linearized inverse problem using the linearized acoustic wave equation is still computational expensive. The convolutional model provides a computationally cheaper way to approximate the linearized acoustic wave equation 2.11. Since high frequency asymptotics and hyperbolic moveout approximation are good approximations only for a layered medium (i.e. the Earth is made of horizontal layers of constant velocity), I need to assume that the acoustic medium is horizontally layered. The principle of original horizontality says that rock layers were originally deposited close to horizontal and continues laterally unless there is a structure or change to prevent its extension [Lyell, 1830]. The layered medium is realistic since much of the Earth's geology consists of successional layers of different rocks, known as "strata".

The horizontally layered medium allows me to write the background velocity as a function of depth

$$v = \hat{v}(z),$$

and write its first order relative perturbation as

$$r = \hat{r}(z).$$

Recalling the medium is horizontally layered, the change of variables

$$z = \int_0^{\frac{t_0}{2}} v(\tau)\, d\tau,$$

allows me to write depth as a function of zero-offset two-way traveltime t_0

$$z = z(t_0).$$

Then I am able to write background velocity as a function of zero-offset two-way traveltime

$$v = \hat{v}(z) = v(t_0),$$

and to write its relative perturbation as a function of zero-offset two-way traveltime as well

$$r = \hat{r}(z) = r(t_0).$$

Notice that in a layered medium with the source and receiver located on the same horizontal plane, there is horizontally translation invariance in the pressure perturbation field. This means that the pressure perturbation field can also be considered as a function of offset (the distance from the source to receiver) or half offset h. I can define another partial linearized forward map

$$F[v]r(t,h) \doteq \hat{F}[v]r(\mathbf{x_r}, t; \mathbf{x_s}),$$

where $h = \frac{|\mathbf{x_r}-\mathbf{x_s}|}{2}$. The most important part of the theory, which is called convolutional model, can be derived based on the high frequency approximation to the linearized acoustic wave equation 2.11 for layered medium. Assuming no noise and primary reflection only, the seismic reflection data can be expressed as the convolutional model for layered medium [Yilmaz, 1987]. As a result of reflection at a single reflector (interface of layers), the source wavelet replicates itself such that it is scaled by the amplitude scaling factor at the reflectors. If there are a number of reflectors, then the source wavelet replicates itself at those reflectors in the same manner.

The convolutional model shows that the predicted seismic reflection data can be approximated by convolving source time function with an amplitude scaled reflectivity. Careful derivation of the convolutional model can be found in Winslow's MA thesis [Winslow, 2000]:

$$F[v]r(t,h) \simeq \omega(t) *_t \{a(T_0(t,h),h)r(T_0(t,h))\}. \tag{2.12}$$

Here t is the two-way traveltime, t_0 is the zero-offset two-way traveltime, h is half offset defined by

$$h = \frac{|\,\mathbf{x_r} - \mathbf{x_s}\,|}{2},$$

$T_0(t,h)$ is a change of variables function, which is the inverse function of the hyperbolic moveout approximation to two-way traveltime $T(t_0,h)$ (to be shown in the next section), $\omega(t)$ is the source time function, $v(t_0)$ is the background velocity, $r(t_0)$ is the relative velocity perturbation with respect to the traveltime at zero offset indicating where are the reflectors, and $a(t_0,h)$ is the amplitude function. The amplitude function can be considered a constant and equal to 1. In the ideal case, I also assume that the point source time function $\omega(t) = \delta(t)$. Since the principle concern of this thesis is kinematic relationship between data and image, the effective replacements of amplitude scaling function by 1 and source time function by a delta function do not seem to invalidate the prediction of the theory. Hence, the simplified convolutional model indicates that predicted seismic reflection data can be approximated by the following formula:

$$F[v]r(t,h) \simeq r(T_0(t,h)). \tag{2.13}$$

This convolutional model is the most important theoretical foundation of DSVA via NMO. Together with hyperbolic moveout approximation which will be discussed next, convolutional model inversion will be derived in the end of this chapter.

2.5 Hyperbolic Moveout Approximation

Traveltime approximations plays an important role in the processing of seismic reflection data. Figure 2.1 shows the simple case of a single horizontal layer. At a given common midpoint M, the reflection traveltime t is to compute the time when waves travel along the ray path from source S_2 to depth point O then back to receiver R_1. Using the Pythagorean theorem, the two-way traveltime t as a function of zero-offset traveltime t_0 and half offset h can be computed from

$$t^2 = t_0^2 + \frac{4h^2}{v^2},$$

where t_0 is twice the traveltime along the vertical path OM, h is half the distance between source and receiver, and v is the velocity of the medium 1 above the reflecting interface (reflector).

Similar to the above simple case, the most employed traveltime approximation is the NMO hyperbola [Dix, 1955]:

$$t^2 \simeq t_0^2 + \frac{4h^2}{v_{\text{RMS}}^2},$$

where RMS (root mean squared) velocity $v_{\text{RMS}}(t_0) = \sqrt{\frac{1}{t_0}\int_0^{t_0} v^2}$. This hyperbolic moveout approximation is a good approximation to the two-way traveltime $T(t_0, h)$

for layered medium and small offset-to-depth ratio. In order to see how to approximate two-way traveltime, I need to start from the Eikonal equation. It has the form

$$\left(\frac{\partial \tau}{\partial x}\right)^2 + \left(\frac{\partial \tau}{\partial z}\right)^2 = \frac{1}{\hat{v}^2(z)}, \tag{2.14}$$

which describes the one-way traveltime propagation from (0,0) to (x, z) in 2D isotropic medium, in which the waves travel at equal speed independently of which direction they travel. In a layered medium with the source and receiver located on the same horizontal plane, there is a horizontally translation invariance in the two-way traveltime. If $\tau_2(x, z)$ is the time of a signal which travels from the source (0,0), reflects at $(\frac{x}{2}, z)$, and back to the receiver $(x, 0)$. The analytical two-way traveltime function

$$T(t_0, h) = \tau_2(x, z),$$

where $x = 2h$ and $z = \int_0^{\frac{t_0}{2}} v$. This two-way traveltime $T(t_0, h) = \tau_2(x, z)$ is related to the solution of Eikonal equation with point source at (0,0) by

$$\tau_2(x, z) = 2\tau(\frac{x}{2}, z).$$

Thus

$$\left(\frac{\partial \tau_2}{\partial x}\right)^2 + \frac{1}{4}\left(\frac{\partial \tau_2}{\partial z}\right)^2 = \frac{1}{\hat{v}^2(z)}.$$

The hyperbolic moveout approximation to the two-way traveltime is just the second order Taylor expansion of $T(t_0, h)$ in h. Details of the continuum derivation of the Dix's hyperbolic moveout approximation can be found from CAAM Technical Report [Symes, 1999]. The following formula represents a good approximation to the two-way

traveltime for layered medium and "small" offset to depth ratio.

$$T(t_0, h) \simeq \sqrt{t_0^2 + \frac{4h^2}{v_{\mathrm{RMS}}^2(t_0)}}, \tag{2.15}$$

$$T_0(T(t_0, h), h) = t_0,$$

where the Root Mean Squared (RMS) velocity is

$$v_{\mathrm{RMS}}(t_0) = \sqrt{\frac{1}{t_0}\int_0^{t_0} v^2}.$$

In order to eliminate confusion with RMS velocity, velocity $v(t_0)$ is called interval velocity in the geophysical literature. The accuracy of the hyperbolic moveout approximation decreases with increasing offset-to-depth ratio. From the approximation shown above, it can be noticed that the two-way traveltime $T(t_0, h)$ explicitly depends on the interval velocity v and the change of variables function $T_0(t, h)$ in 2.13 is the inverse function of the two-way traveltime function $T(t_0, h)$.

2.6 Convolutional Model Inversion

Combining the convolutional model 2.13 with this hyperbolic moveout approximation, the convolutional model inversion is:

Given observed seismic reflection data d, find smooth background velocity v so that

$$r(T_0(t, h)) \simeq d(t, h).$$

Applying a velocity-dependent change of variables $T(t_0, h)$, it maps seismic reflection data $d(t, h)$ from $t - h$ domain to $t_0 - h$ domain and forms redundant seismic images

(NMO corrected traces) defined by $d(T(t_0, h), h)$. This procedure is called NMO correction in the geophysical literatures, which maps events (consistent lining ups of several wiggles over several traces) in the seismic reflection data to where they come from. One NMO corrected trace corresponds to one offset. The NMO corrected traces can be considered as zero-offset recording of reflection data, which means that the sources are placed at each h and the primary reflection signals are received at the same locations as their sources. Since it is assumed that the medium is horizontally layered, the NMO corrected traces using the accurate velocity should be horizontally invariant. It means that the NMO corrected traces using the accurate velocity should be the same, independent of offset. Hence, consistency of redundant NMO corrected traces can be evaluated by the flatness or alignment of same events between different traces.

Chapter 3

Methods

This chapter introduces the motivation for differential semblance which opens up a new approach to velocity analysis, reviews previous approaches to differential semblance, and analyzes the advantages and disadvantages of the previous implementation. DS-CR is shown in detail following the review. Four fundamental stages are involved in the last section of implementation: data preparation; velocity representation; numerical computation; numerical optimization.

3.1 Motivation for Differential Semblance

Some researchers considered velocity analysis, the process of estimating seismic velocity, as an optimization problem in the form of a misfit function to be minimized [Kolb et al., 1986]. The least squares method is among the most widely investigated one. The least squares approach is to find the optimal velocity such that the sum of

squared residuals, which is the difference between observed and predicted data, has its least value. However this method may fail if the starting velocity model is far enough from the true velocity model, because of the existence of the secondary solutions [Kolb et al., 1986] [Gauthier et al., 1986]. If a gradient method is used to solve the optimization problem, the iteration process may stop at a local minimum that may be far from the global minimum. To avoid this, global optimization methods such as simulated annealing or genetic algorithms can be used [Kirkpatrick et al., 1983] [Sen and Stoffa, 1991b] [Sen and Stoffa, 1991a]. However, the associated computational cost of global search methods is too high to be suitable for industrial scale velocity estimation.

Another strategy of velocity analysis is migration velocity analysis as mentioned in the chapter 1. The idea of migration velocity analysis is that redundant images are consistent when the velocity is correct. Most of the objective functions used in the literature of migration velocity analysis are called semblance [Neidell and Taner, 1971]. Semblance is a coherency measurement for the recorded seismic data. The data can be sorted by common midpoint (CMP), which is is the point on the surface halfway between the source and the receiver. A CMP gather is a gather of seismic traces that share the same CMP. The idea of classical semblance was first introduced by Toldi [Toldi, 1989]. The Classical Semblance (CS) function provides normalized measurement of the stack power of CMP gathers and defined as

$$J_{CS}[v] = \frac{\sum_{t_0}(\sum_h d(T(t_0,h),h))^2}{\sum_{t_0}\sum_h d(T(t_0,h),h)^2}. \tag{3.1}$$

where h is the half offset, t_0 is zero-offset traveltime, $T(t_0, h)$ is the two-way traveltime function. The sums over all t_0 and h in the above discrete form of CS objective function are approximations of the numerical integration using the composite trapezoidal rule in order to get the L^2 norms in the continuous form of the CS objective function. As explained in section 2.6, $d(T(t_0, h), h)$ is the image after NMO correction which can be tested for consistency by evaluating the alignments of the same event between different traces in zero-offset traveltime and half offset domain for layered medium. $\sum_h d(T(t_0, h), h)$ represents that the traces are added together (stack) to form a new single stacked trace. The stack power is to the sum of the squared value along the new single stacked trace, written as

$$J_{SP}[v] = \sum_{t_0} (\sum_{h} d(T(t_0, h), h))^2, \tag{3.2}$$

The stack power is one way to evaluate the alignment since the maximum stack power can be achieved if the velocity is correct. The classical semblance shown in 3.1, which is normalized measurement of the stack power, has the same behavior as stack power shown in 3.2. Unfortunately, Chauris and Noble showed the examples that the classical semblance has similar defect as least squares [Chauris and Noble, 2001]. A discussion, which shows that Classical Semblance is equivalent to least squares, can be found in the Appendix A. Classical Semblance is highly non-convex with many local maxima and cannot be applied with a local optimization process.

An alternative is differential semblance velocity analysis (DSVA), proposed by Symes [Symes and Carazzone, 1991] [Symes, 1999], which opens up a new approach

to velocity analysis due to its global convexity and asymptotic stability [Symes, 1999]. DSVA is able to provide good results even when the initial velocity model is not close to the true velocity model. Considered as a convex optimization problem, DSVA can be carried out by means of a gradient-based method, thereby avoiding a costly global search method. For migration velocity analysis, DSVA is the only approach so far that allows automatic velocity computation by means of local optimization process.

DSVA via NMO correction is the simplest form of DS and can be defined as

$$J_{DS}[v] = \sum_{t_0} \sum_{h} \left(\frac{\partial}{\partial h} d(T(t_0, h), h) \right)^2 . \tag{3.3}$$

As explained in section 2.6, the NMO corrected traces using the accurate velocity should be horizontally invariant and DSVA via NMO correction evaluates the consistency of redundant NMO corrected traces by the flatness of the same event between different traces. It is easy to see from the above formula that DS measures the flatness of the same events by computing derivatives with respect to half offset. The comparison between CS and DS in chapter 4 will show that DS exhibits much better properties in terms of global convexity than CS as illustrated in Figure 5.9 and 5.10.

3.2 Approaches to Differential Semblance

There are many different approaches to DS which can be divided as two classes. The first class is DS variants based on surface oriented extension. Surface oriented extension is usually carried out via a ray-based method. Ray-based methods consist

of wide-band rays (based on high frequency asymptotics) which are traced by solving the traveltime equation. The other one is variants based on depth oriented extension. Depth oriented extension is wave-based method. Wave-based methods consist of band-limited wavefields which are constructed by solving the wave equation for all depth. Since the method I developed is based on surface oriented extension, I only discuss DS variants based on surface oriented extension in this section.

Convolutional model simulation of plane wave data

The early approach to DS has been formulated in detail for the convolutional approximation to the layered constant-density acoustic model and was appropriate to plane-wave seismograms [Symes and Carazzone, 1991] [Minkoff and Symes, 1997]. The image in $\tau - p$ domain (intercept time - slowness domain) was transformed from $t - x$ domain (time - offset domain) through a Radon transform. This $\tau - p$ section was flattened by updating the velocity automatically. They compared semblance between neighboring traces and measured the deviation from flatness. Theoretical and numerical analysis of this version of DS indicated that stable and reasonably accurate estimates of velocity can be derived.

Convolutional model simulation of common midpoint data

Symes concluded that a local gradient based optimization method can be used to minimize the DS objective function based on convolutional model typically by using com-

mon midpoint data [Symes, 1993] [Symes and Gockenbach, 1995] [Symes, 1998b]. A recent implementation of this method was developed by Li and Symes [Li and Symes, 2007]. Their approach, using hyperbolic normal moveout correction, is effective within the limits that the medium is mild lateral heterogeneous (velocity does not variate much in the horizontal direction) and the data are dominated by primary events. They also combined DS via NMO correction with the multiple filter suggested in Mulder's paper [Mulder and ten Kroode, 2002] to decrease the effect of multiple reflections on the velocity estimate. This NMO-driven DS algorithm is very simple and fast relative to other variants based on more faithful approximations to wave physics, but nonetheless is accurate enough to be used to process seismic data with some success. Therefore I focus on this simplest form of DS in this thesis.

Generalized Radon transform simulation of acoustic scattering

Chauris and Noble demonstrated the efficiency and accuracy of the DS method for velocity model estimation which simulated acoustic scattering by generalized Radon transform [Chauris and Noble, 2001]. While convolutional model is based on layered media, generalized Radon transform simulation can deal with laterally heterogeneous media. The linearized description of seismic scattering which they used approximates only primary reflection. It neglects multiple reflections, so that multiple energy in the data appears as coherent noise. Thus their success has a prerequisite that a careful pre-processing has to be applied before the inversion in order to eliminate multiple

and other coherent noise. But they did not propose any effective way to suppress the multiple reflections apart from the classical pre-processing.

Mulder and ten Kroode suggested a new way to deal with these multiple reflections [Mulder and ten Kroode, 2002]. Their objective function mimics Symes's differential semblance objective function [Symes, 1998b]. Their method was able to handle the multiples effectively by incorporating a multiple filter inside the differential semblance objective function.

Generalized Radon transform simulation of anisotropic elastic scattering

The DS variants in the preceding paragraphs showed that the DS methods are able to deal with isotropic acoustic scattering. In this paragraph, I want to show that DS can be used as an approach to simulate the anisotropic elastic scattering as well. This approach involved an extension of DS to converted waves [de Hoop et al., 2005]. When a P (primary) wavefront from a source strikes an interface in the elastic scattering, the reflected energy is partitioned into P (primary) and S (shear) waves. This approach demonstrated the use of MVA on joint PP and PS reflection seismic data while the DS variants shown in the previous paragraphs only dealt with PP reflection seismic data. This method was based on the DS objective function parameterized by scattering angle and azimuth to find fitting velocity model and showed the potential of DS to deal with multiple wave speeds.

3.3 A New Formula of Differential Semblance

Recall the convolutional model inversion shown in section 2.6:

given observed seismic reflection data d, find smooth background velocity v so that

$$r(T_0(t,h)) \simeq d(t,h).$$

I introduce the nonphysical reflectivity model $r(t_0,h)$ by extending the definition of reflectivity to depend on more spatial degree of freedom. Convolutional model can be recovered if and only if the extra degree of freedom is present in some specific way, satisfying the constraint $\frac{\partial r}{\partial h} = 0$. The inverse problem is

$$\min_{v,r} \quad J[v,r] = \frac{1}{2}\|r - d \circ T\|_2^2 \tag{3.4}$$

$$s.t. \quad \left\|\frac{\partial r}{\partial h}\right\|_2^2 = 0 \tag{3.5}$$

The operator $d \circ T$ is defined as $d \circ T(t_0,h) = d(T(t_0,h),h)$, which is the image after NMO correction. The objective function of this optimization problem is easy to minimize without constraint. The optimal solution is $[v,r]$ such that for any background velocity model v, let the nonphysical reflectivity model r be the NMO corrected image

$$r[v](t_0,h) = d(T[v](t_0,h),h).$$

Then for noise-free data, the optimization problem 3.4 and 3.5 has same solution as

the following problem

$$\min_{v,r} \quad J[v,r] = \frac{1}{2}\left\|\frac{\partial r}{\partial h}\right\|_2^2 \tag{3.6}$$

$$s.t. \quad \|r - d \circ T\|_2^2 = 0 \tag{3.7}$$

Notice that for noisy data, those two problem are different with different solutions. Therefore the NMO-based DS objective function becomes

$$J[v] = \frac{1}{2}\left\|\frac{\partial}{\partial h} d \circ T\right\|_2^2, \tag{3.8}$$

when the constraint 3.7 is eliminated.

The NMO-based DS objective function in the previous implementation for a single CMP gather is

$$J_{DS-CV}[v] = \frac{1}{2}\int dh \int dt_0 \frac{\partial T}{\partial t_0}(t_0,h)\left|\frac{\partial}{\partial h} d(T(t_0,h),h)\right|^2. \tag{3.9}$$

This objective function was optimized by the DS-CV approach [Li and Symes, 2007]. The factor $\frac{\partial T}{\partial t_0}$ is to make the gradient computation easier, not effecting the stationary points since it has to be always greater than 0 (see 3.12). This means that the global minimizer of the DS-CV objective function is also the optimal velocity that minimizes the general DS objective function (3.8). DS-CV was accurate enough to be used to process seismic data with some success. But the seismic data $d(t,h)$ is not sampled on the point $(T(t_0,h),h)$ so that this approach involved the interpolation of the oscillatory seismic data. Spurious local minima of the DS-CV objective function are shown in my numerical experiments due to the large interpolation error. In order to

avoid the interpolation of the oscillatory seismic data, I employs a new discretization method (DS-CR) to approximate the same continuous DS objective function as DS-CV by using the chain rule, which is the first advantage of DS-CR over DS-CV.

At first, I apply the chain rule,

$$\begin{aligned}\frac{\partial}{\partial h}d(T(t_0,h),h) &= \left(\frac{\partial d}{\partial h}+\frac{\partial T}{\partial h}\frac{\partial d}{\partial t}\right)(T(t_0,h),h)\\ &= \left(\frac{\partial d}{\partial h}+p\frac{\partial d}{\partial t}\right)(t,h).\end{aligned}$$

Here p is the ray parameter or slowness, defined by

$$p(t,h)=\frac{\partial T}{\partial h}(t_0,h)|_{t_0=T_0(t,h)}.$$

The two-way traveltime function $T(t_0,h)$ is computed by

$$T(t_0,h)=\sqrt{t_0^2+4u(t_0)h^2},$$

where RMS square slowness $u(t_0)=\frac{t_0}{\int_0^{t_0}v^2}$ under the assumption that the offset is small compared to depth. Computing the partial derivative

$$\frac{\partial T}{\partial h}(t_0,h)=\frac{4h}{T(t_0,h)}u(t_0)$$

leads to

$$p(t,h)=\frac{\partial T}{\partial h}(t_0,h)|_{t_0=T_0(t,h)}=\frac{4h}{t}u(T_0(t,h)).$$

The DS-CR objective function, equivalent to the DS-CV objective function $J_{DS-CV}[v]$ with an alternative form in the continuous sense, is written as

$$J_{DS-CR}[v]=\frac{1}{2}\int dh\int dt\ \left|\left(p\frac{\partial d}{\partial t}+\frac{\partial d}{\partial h}\right)(t,h)\right|^2, \tag{3.10}$$

where the slowness is $p(t,h) = \frac{4h}{t}u(T_0(t,h))$ and RMS square slowness is $u(t_0) = \frac{t_0}{\int_0^{t_0} v^2}$.

Since the calculation takes the same form for all CMP gathers, I extend the dependence on midpoint and consider all the CMP gathers in the final result.

Suppose shot and receiver coordinates are denoted x_s, y_s and x_r, y_r respectively and they are located on the surface. Data trace $d(t,h,x,y)$ are sorted by CMP coordinates

$$x = \frac{x_r + x_s}{2}, y = \frac{y_r + y_s}{2}.$$

Here t is the two-way traveltime and h is the half offset defined by

$$h = \frac{\sqrt{(x_r - x_s)^2 + (y_r - y_s)^2}}{2}.$$

The RMS square slowness $u(t_0,x,y)$ is defined by

$$u(t_0,x,y) = \frac{t_0}{\int_0^{t_0} v(\tau,x,y)^2\, d\tau}.$$

The hyperbolic traveltime approximation is

$$T(t_0,h,x,y) = \sqrt{t_0^2 + 4u(t_0,x,y)h^2}. \tag{3.11}$$

The function $T_0(t,h,x,y)$ is the inverse function of $T(t_0,h,x,y)$. The function $T_0(t,h,x,y)$ is well defined where

$$0 < \frac{\partial T}{\partial t_0}(t_0,h,x,y) = \frac{t_0 + 2\frac{\partial u}{\partial t_0}(t_0,x,y)h^2}{T(t_0,h,x,y)} \tag{3.12}$$

[Symes, 1999]. The interval velocities $v(t_0,x,y)$ are bounded by the velocity envelopes $v_{max}(t_0,x,y)$ and $v_{min}(t_0,x,y)$.

With all the CMP gathers, the objective function is

$$J_{DS-CR}[v] = \frac{1}{2}\int dx \int dy \int dh \int_0^{t_{max}} dt \left| \left(p\frac{\partial d}{\partial t} + \frac{\partial d}{\partial h} \right)(t,h,x,y) \right|^2 . \tag{3.13}$$

Since $J_{DS-CR}[v]$ has no spurious local minimizers, I use a Newton-like, gradient-based optimization method to minimize the DS-CR objective function. It is necessary to compute the gradient of J_{DS-CR} with respect to v. The expression for the gradient is

$$\begin{aligned} \nabla J_{DS-CR}[v](\tau,x,y) = -\int dh\ v(\tau,x,y) \int_{T(\tau,h,x,y)}^{t_{max}} dt \\ \left\{ \frac{\partial T_0}{\partial t}(t,h,x,y) \frac{p^2(t,h,x,y)}{2h} \frac{\partial d}{\partial t}(t,h,x,y) \left(p\frac{\partial d}{\partial t} + \frac{\partial d}{\partial h} \right)(t,h,x,y) \right\}. \end{aligned} \tag{3.14}$$

A derivation of this formula is given in the Appendix B. Another simpler way to derive this gradient formula is based on the DS-CV gradient [Li and Symes, 2007]

$$\begin{aligned} \nabla J_{DS-CV}[v](\tau,x,y) = -\int dh\ v(\tau,x,y) \int_{\tau}^{t_{0max}} dt_0 \\ \left\{ u^2(\tau,x,y) \frac{8h}{T(t_0,h,x,y)^2} \left(\frac{\partial d}{\partial t}(T(t_0,h,x,y),h,x,y) \right) \left(\frac{\partial}{\partial h} d(T(t_0,h,x,y),h,x,y) \right) \right\}, \end{aligned} \tag{3.15}$$

and apply change of variables $t_0 \rightarrow T_0(t,h)$, which is the inverse of the two-way traveltime function. After comparing the gradients of DS-CV (3.15) and DS-CR (3.14), it can be noticed that DS-CV involves interpolation of oscillatory integrand, which affects every term in the integral. However, DS-CR only involve interpolation of oscillatory integrand on the lower boundary terms in the integral with respect to t. Since there are thousands of time samples, DS-CR should have much less interpolation error in the computation of gradient as well, which is another advantage of DS-CR over DS-CV.

3.4 Implementation

The implementation is done with the C++/RVL framework [Padula et al., 2004] and SU package [Stockwell and Cohen, 2004]. RVL is a mathematical library which provides simple access to some very powerful optimization algorithms which helps me optimize the objective function. SU is a powerful tool to generate, analyze and process seismic data which I use in this thesis.

Referring to equations 3.11, 3.13 and 3.14, the implementation involves four fundamental stages: data preparation; velocity representation; numerical computation of the integrals, derivatives, and interpolations; numerical optimization.

3.4.1 Data Format and Velocity Representation

The SU data format is based on the `SEG-Y` format [Barry and Kneale, 1980]. The SU format consists of data traces each of which has a header. Both the header and the trace data are written in the binary format. The SU header contains every attributes I need to analyze and process the data such as `offset`(offset), `dt`(time sampling rate), `sx`(X source coordinate), `sy`(Y source coordinate), `gx`(X receiver coordinate), `gy`(Y receiver coordinate), and etc. The data to be processed is sorted by CMP gathers and can be read and written by the SU package. SU commands can be also used to generate the synthetic seismogram for linear velocity function by using `susynlv` and to view the seismic data set, to apply a zero-phase, sine-squared tapered filter, or to view the frequency spectrum of the data by using `suximage, sufilter, suspecfx,...`

The data in equation 3.13 and 3.14 is gridded at the level which is adequate to represent the oscillatory seismic data. Theory developed in the preceding chapter presumes that the velocity must be smooth on the wavelength scale. The success of iterative methods for my differential semblance might depend on the velocity encountered during the iteration remaining similarly smooth. Hence, the gridding for the oscillatory data gives too many degrees of freedom for the velocity. To control the smoothness of velocity updates, it is necessary to reduce the number of degrees of freedom for velocity so that the method is effective. In order to reduce the number of degrees of freedom in the velocity model, I employ a fixed nodal structure for velocity representation to produce relatively smooth velocity model. This nodal placement is static, nodal values should be bounded by the nodal values of upper and lower envelope velocities. The simplest velocity representation meet those requirements is the piecewise linear interpolation. The velocity data structure is obtained by piecewise linear interpolation of the nodal values and sampled on the (t_0, x, y) grid.

3.4.2 Numerical Computation

The numerical computation in this thesis involves computation of the integrals, derivatives, and interpolations. The integrals in equations 3.11, 3.13 and 3.14 can be replaced by composite trapezoidal rule and have second-order accuracy. The detail of the cubic interpolation and the computation of the derivative are given below.

3.4.2.1 Computation of t_0 and p

For a single CMP gather, the traveltime table $T(t_0, h)$ is an important tool for doing the change of variables of t to t_0 conversion. This mapping is computed using a numerical scheme that is governed by equation 2.15. Since the RMS square slowness is defined as

$$u(t_0) = \frac{t_0}{\int_0^{t_0} v(t_0)^2},$$

for a single CMP gather, $u_k = u(t_0[k])$ for $k = 1, ..., n$ can be computed using the composite trapezoid rule:

$$u_k \simeq \frac{t_0[k]}{\sum_{i=0}^{k-1} \Delta t \times v(t_0[i])^2},$$

for $k = 1, ..., n$. Here Δt is the time sampling rate, ot_0 is the first point of t_0, considered as 0th sample, and $t_0[k]$ is the kth sample of t_0 which can be computed by $t_0[k] = ot_0 + k\Delta t$.

Time traces are discretized and have $n+1$ samples from ot to t_{max} with sample rate Δt. The ith sample occurs at $t[i] = ot + i\Delta t$ for $i = 0, 1, ..., n$. Two-way traveltime t has the same discretization as two-way traveltime at zero offset t_0. But $T(t_0, h)$ is very unlikely to coincide with one of the sample time $t[i]$. Thus the implementation of the RMS square slowness involving $T_0(t, h)$ inevitably requires interpolation, which is not smooth with respect to velocity parametrization. In order to minimize this difficulty, I adopt the local cubic Lagrangian interpolation to compute the RMS square slowness.

Define $u_k = u(t_0[k])$ for $k = 1, ..., n$ and $U_i = u(T_0(t[i], h))$ for $i = m, ..., n$ where $m = \lceil \frac{T(0,h)}{\Delta t} \rceil$.

How to compute U_i which is u at time $t = t[i]$?

First finding out the index k of t_0 for which $T(t_0[k-1], h) < t < T(t_0[k], h)$ using the traveltime table, I can compute $u_{k-1}, u_k, u_{k+1}, u_{k+2}$.

Let $t_c = T(t_0[k], h)$, $t_{-1} = t_c - \Delta t$, $t_1 = t_c + \Delta t$, and $t_2 = t_c + 2\Delta t$.

I obtain U_i using the cubic Lagrangian interpolation:

$$l_{-1} = \frac{(t-t_c)(t-t_1)(t-t_2)}{(t_{-1}-t_c)(t_{-1}-t_1)(t_{-1}-t_2)}$$

$$l_0 = \frac{(t-t_{-1})(t-t_1)(t-t_2)}{(t_c-t_{-1})(t_c-t_1)(t_c-t_2)}$$

$$l_1 = \frac{(t-t_{-1})(t-t_c)(t-t_2)}{(t_1-t_{-1})(t_1-t_c)(t_1-t_2)}$$

$$l_2 = \frac{(t-t_{-1})(t-t_c)(t-t_1)}{(t_2-t_{-1})(t_2-t_c)(t_2-t_1)}$$

$$U_i \simeq l_{-1}u_{k-1} + l_0 u_k + l_1 u_{k+1} + l_2 u_{k+2}.$$

Hence slowness $p_i = p(t[i], h)$ can be obtained by

$$p_i = \frac{4h}{t[i]} U_i,$$

for $i = m, .., n$. This Lagrangian interpolation has fourth order accuracy.

3.4.2.2 Computation of $\frac{\partial d}{\partial t}$ and $\frac{\partial d}{\partial h}$

I use cell centered difference to approximate $\frac{\partial d}{\partial t}$ and $\frac{\partial d}{\partial h}$. Since the data are ordered by increasing offset within each CMP gather, this cell difference requires only one additional trace buffer.

Write $t_i = ot + i\Delta t$, $h_j = oh + j\Delta h$, and $d_{ij} = d(t_i, h_j)$. Centered difference approximation can be expressed by

$$\frac{\partial d}{\partial t}(t_{i+\frac{1}{2}}, h_{j+\frac{1}{2}}) \simeq \frac{1}{2} \times (\frac{d_{i+1,j} - d_{i,j}}{\Delta t} + \frac{d_{i+1,j+1} - d_{i,j+1}}{\Delta t}),$$

$$\frac{\partial d}{\partial h}(t_{i+\frac{1}{2}}, h_{j+\frac{1}{2}}) \simeq \frac{1}{2} \times (\frac{d_{i,j+1} - d_{i,j}}{\Delta h} + \frac{d_{i+1,j+1} - d_{i+1,j}}{\Delta h}).$$

This is a second-order accurate method.

3.4.3 Error Estimation

One advantage of using DS-CR over DS-CV is that the DS-CR objective function only involves interpolation of smooth functions instead of interpolation of oscillatory seismic data. The error in the computation of the DS-CR objective function is mainly created by approximation to the partial derivatives. This section provides sufficient conditions for a good approximation to the partial derivatives, furthermore to the objective function. The Fourier expansion of the seismic data in the frequency domain with single reflector at zero offset traveltime t_0 can be written as,

$$\hat{d}(\omega, h) = \sum_{n=-\infty}^{\infty} C_n e^{-i\omega T(t_0, h)},$$

where C_n are the corresponding Fourier coefficients, ω is the angular frequency, h is the half offset, $T(t_0, h)$ is the two-way traveltime function, and $\hat{d}(\omega, h)$ is the seismic data. For simplification, I only consider single angular frequency ω, single reflector at zero offset traveltime t_0, amplitude A, and true velocity model v. The data in the

traveltime domain (I only consider the real part) can be written as

$$d(t,h) = A\sin\omega\left(t - \sqrt{t_0^2 + 4h^2u(t_0)}\right),$$

where t is the traveltime, h is the half offset, and RMS square slowness $u(t_0) = \frac{t_0}{\int_0^{t_0} v(\tau)^2\, d\tau}$.

Let $d_1 = d_{i,j+\frac{1}{2}} = d(t_i, h_j + \frac{1}{2}\Delta h)$, $d_2 = d_{i+1,j+\frac{1}{2}} = d(t_i + \Delta t, h_j + \frac{1}{2}\Delta h)$, $d_{\xi_1} = d(t_i, \xi_1)$ for some $\xi_1 \in (h_j, h_{j+\frac{1}{2}})$, $d_{\xi_2} = d(t_i, \xi_2)$ for some $\xi_2 \in (h_{j+\frac{1}{2}}, h_{j+1})$, $d_{\xi_3} = d(t_{i+1}, \xi_3)$ for some $\xi_3 \in (h_j, h_{j+\frac{1}{2}})$, and $d_{\xi_4} = d(t_{i+1}, \xi_4)$ for some $\xi_4 \in (h_{j+\frac{1}{2}}, h_{j+1})$.

$$d_{i,j} = d_1 - \frac{\Delta h}{2}\frac{\partial d_1}{\partial h} + \frac{1}{2}\left(\frac{\Delta h}{2}\right)^2\frac{\partial^2 d_1}{\partial h^2} - \frac{1}{6}\left(\frac{\Delta h}{2}\right)^3\frac{\partial^3 d_{\xi_1}}{\partial h^3} \tag{3.16}$$

$$d_{i,j+1} = d_1 + \frac{\Delta h}{2}\frac{\partial d_1}{\partial h} + \frac{1}{2}\left(\frac{\Delta h}{2}\right)^2\frac{\partial^2 d_1}{\partial h^2} + \frac{1}{6}\left(\frac{\Delta h}{2}\right)^3\frac{\partial^3 d_{\xi_2}}{\partial h^3} \tag{3.17}$$

$$d_{i+1,j} = d_2 - \frac{\Delta h}{2}\frac{\partial d_2}{\partial h} + \frac{1}{2}\left(\frac{\Delta h}{2}\right)^2\frac{\partial^2 d_2}{\partial h^2} - \frac{1}{6}\left(\frac{\Delta h}{2}\right)^3\frac{\partial^3 d_{\xi_3}}{\partial h^3} \tag{3.18}$$

$$d_{i+1,j+1} = d_2 + \frac{\Delta h}{2}\frac{\partial d_2}{\partial h} + \frac{1}{2}\left(\frac{\Delta h}{2}\right)^2\frac{\partial^2 d_2}{\partial h^2} + \frac{1}{6}\left(\frac{\Delta h}{2}\right)^3\frac{\partial^3 d_{\xi_4}}{\partial h^3} \tag{3.19}$$

The centered difference approximation to the h partial derivative is

$$\frac{\partial d}{\partial h}(t_{i+\frac{1}{2}}, h_{j+\frac{1}{2}}) \simeq \frac{1}{2}\times\left(\frac{d_{i,j+1} - di,j}{\Delta h} + \frac{d_{i+1,j+1} - d_{i+1,j}}{\Delta h}\right) \triangleq \left(\frac{\partial d}{\partial h}\right)_N.$$

Using the continuous property of the data with respect to h, compute the numerical approximation to the h derivative

$$\begin{aligned}\left(\frac{\partial d}{\partial h}\right)_N &= \frac{\frac{\partial d_1}{\partial h} + \frac{\partial d_2}{\partial h}}{2} + \frac{\Delta h^2}{24}\frac{\frac{\partial^3 d_{\xi 5}}{\partial h^3} + \frac{\partial^3 d_{\xi 6}}{\partial h^3}}{2} \\ &= \frac{\partial d}{\partial h}(t_{i+\frac{1}{2}}, h_{j+\frac{1}{2}}) + \frac{\Delta t^2}{8}\frac{\partial}{\partial t^2}\left(\frac{\partial d_{\xi 7}}{\partial h}\right) + \frac{\Delta h^2}{24}\frac{\partial^3 d_{\xi 8}}{\partial h^3},\end{aligned}$$

where $d_{\xi_5} = d(t_i, \xi_5)$ for some $\xi_5 \in (h_j, h_{j+1})$, $d_{\xi_6} = d(t_{i+1}, \xi_6)$ for some $\xi_6 \in (h_i, h_{j+1})$, $d_{\xi_7} = d(\xi_7, h_{j+\frac{1}{2}})$ for some $\xi_7 \in (t_i, t_{i+1})$, and $d_{\xi_8} = d(t_{i+\frac{1}{2}}, \xi_8)$ for some $\xi_8 \in (h_j, h_{j+1})$. Then the approximation error is

$$\left| \left(\frac{\partial d}{\partial h} \right)_N - \frac{\partial d}{\partial h} \right| = \left| \frac{\Delta h^2}{24} \frac{\partial^3 d_{\xi_8}}{\partial h^3} + \frac{\Delta t^2}{8} \frac{\partial}{\partial t^2} \left(\frac{\partial d_{\xi_7}}{\partial h} \right) \right|,$$

where $\left(\frac{\partial d}{\partial h}\right)_N$ denotes the numerical approximation to $\frac{\partial d}{\partial h}$. Since

$$d(t, h) = A \sin \omega (t - \sqrt{t_0^2 + 4h^2 u}),$$

then

$$\frac{\partial d}{\partial h}(t, h) = -\frac{4uA\omega h \cos \omega \left(t - \sqrt{t_0^2 + 4h^2 u(t_0)} \right)}{\sqrt{t_0^2 + 4h^2 u(t_0)}}.$$

$$\begin{aligned} \frac{\partial^2 d}{\partial h^2}(t, h) &= 4uA\omega \sin \omega \left(t - \sqrt{t_0^2 + 4h^2 u} \right) \frac{-4h^2 u\omega}{t_0^2 + 4h^2 u} - \\ &\quad -4uA\omega \cos \omega \left(t - \sqrt{t_0^2 + 4h^2 u} \right) \frac{\sqrt{t_0^2 + 4h^2 u} - h \frac{\partial}{\partial h} \sqrt{t_0^2 + 4h^2 u}}{t_0^2 + 4h^2 u} \\ &= 4uA\omega \sin \omega \left(t - \sqrt{t_0^2 + 4h^2 u} \right) \frac{-4h^2 u\omega}{t_0^2 + 4h^2 u} - \\ &\quad -4uA\omega \cos \omega \left(t - \sqrt{t_0^2 + 4h^2 u} \right) \frac{t_0^2}{(t_0^2 + 4h^2 u)^{\frac{3}{2}}}. \end{aligned}$$

Write $S = \sin \omega \left(t - \sqrt{t_0^2 + 4h^2 u} \right)$, $C = \cos \omega \left(t - \sqrt{t_0^2 + 4h^2 u} \right)$, and $T = T(t_0, h) = \sqrt{t_0^2 + 4h^2 u}$.

$$\frac{\partial^2 d}{\partial h^2}(t, h) = -(4u\omega)^2 A \frac{Sh^2}{T^2} - 4u\omega A t_0^2 \frac{C}{T^3}.$$

$$\frac{\partial^3 d}{\partial h^3}(t, h) = -(4u\omega)^2 A \left(-C \frac{4h^3 u\omega}{T^3} + S \frac{2ht_0^2}{T^4} \right) - 4u\omega A t_0^2 \left(S \frac{4hu\omega}{T^4} - 3C \frac{4hu}{T^5} \right).$$

$$\frac{\partial^2}{\partial t^2} \left(\frac{\partial d}{\partial h} \right)(t, h) = -\frac{4u\omega^3 AhC}{T}.$$

Notice that

$$\frac{\partial d}{\partial h}(t,h) = -\frac{4u\omega AhC}{T}.$$

The relative error bound is

$$\left|\frac{\left(\frac{\partial d}{\partial h}\right)_N - \frac{\partial d}{\partial h}}{\frac{\partial d}{\partial h}}\right| \leq \max_{\Delta h,h,\Delta t,\omega} \left|\frac{\Delta h^2}{24}\left(-\frac{16h^2u^2\omega^2}{T^2} + \frac{12u\omega t_0^2}{T^3}\frac{S}{C} - \frac{12ut_0^2}{T^4}\right) + \frac{\Delta t^2}{8}\omega^2\right|.$$

Consider $t \sim \sqrt{t_0^2 + 4h^2u}$ and angular frequency ω is actually 2π times frequency f, I can get sufficient conditions for a good centered difference approximation to h:

$$\frac{\Delta t^2}{2}\pi^2 f^2 \leq \epsilon \tag{3.20}$$

$$\frac{\Delta h^2}{6}\left(\frac{16\pi^2h^2u^2f^2}{T^2} + \frac{6\pi t_0^2uf}{T^3} + \frac{3t_0^2u}{T^4}\right) \leq \epsilon. \tag{3.21}$$

Similarly, the centered difference approximation to t derivatives is

$$\frac{\partial d}{\partial t}(t_{i+\frac{1}{2}}, h_{j+\frac{1}{2}}) \simeq \frac{1}{2} \times \left(\frac{d_{i+1,j} - di,j}{\Delta t} + \frac{d_{i+1,j+1} - d_{i,j+1}}{\Delta t}\right).$$

$$\begin{aligned}
\frac{\partial d}{\partial t} &= A\omega C. \\
\frac{\partial^2 d}{\partial t^2} &= -A\omega^2 S. \\
\frac{\partial^3 d}{\partial t^3} &= -A\omega^3 C. \\
\frac{\partial^2}{\partial h^2}\left(\frac{\partial d}{\partial t}\right) &= 4Au\omega^2\left(-C\frac{4h^2u\omega}{T^2} + S\frac{t_0^2}{T^3}\right).
\end{aligned}$$

The relative error bound is

$$\left|\frac{\left(\frac{\partial d}{\partial t}\right)_N - \frac{\partial d}{\partial t}}{\frac{\partial d}{\partial t}}\right| \leq \max_{\Delta h,h,\Delta t,\omega} \left|\frac{\Delta h^2}{8}\left(-\frac{16h^2u^2\omega^2}{T^2} + \frac{4u\omega t_0^2}{T^3}\frac{S}{C}\right) - \frac{\Delta t^2}{24}\omega^2\right|.$$

Constraints 3.20 and 3.21 are also sufficient conditions for a good centered difference approximation to t. Now suppose I have

$$\left| \frac{\left(\frac{\partial d}{\partial h}\right)_N - \frac{\partial d}{\partial h}}{\frac{\partial d}{\partial h}} \right| \leq \epsilon, \tag{3.22}$$

$$\left| \frac{\left(\frac{\partial d}{\partial t}\right)_N - \frac{\partial d}{\partial t}}{\frac{\partial d}{\partial t}} \right| \leq \frac{\epsilon}{2}, \tag{3.23}$$

and fouth order Lagrangian interpolation allows relative slowness error $\left|\frac{p_N - p}{p}\right| \leq \frac{\epsilon}{4}$. Next I want to compute the numerical approximation error of $\left(p\frac{\partial d}{\partial t}\right)_N$, where p is the slowness and P_N is the numerical approximation to p.

$$\frac{\left(p\frac{\partial d}{\partial t}\right)_N - p\frac{\partial d}{\partial t}}{p\frac{\partial d}{\partial t}} = \frac{p_N\left(\frac{\partial d}{\partial t}\right)_N - p\left(\frac{\partial d}{\partial t}\right)_N}{p\frac{\partial d}{\partial t}} + \frac{p\left(\frac{\partial d}{\partial t}\right)_N - p\frac{\partial d}{\partial t}}{p\frac{\partial d}{\partial t}}$$

The absolute value of the second term is smaller than $\frac{\epsilon}{2}$ by 3.23. First term is

$$\left(\frac{p_N - p}{p}\right) \times \left(\frac{\left(\frac{\partial d}{\partial t}\right)_N}{\frac{\partial d}{\partial t}}\right).$$

From 3.23, it is known that

$$\left| \frac{\left(\frac{\partial d}{\partial t}\right)_N}{\frac{\partial d}{\partial t}} \right| \leq 1 + \frac{\epsilon}{2}.$$

Since $\left|\frac{p_N - p}{p}\right| \leq \frac{\epsilon}{4}$,

$$\left| \frac{\left(p\frac{\partial d}{\partial t}\right)_N - p\frac{\partial d}{\partial t}}{p\frac{\partial d}{\partial t}} \right| \leq \epsilon.$$

Denote $I_{ij} = I(t_i, h_j) = \left(\frac{\partial d}{\partial h} + p\frac{\partial d}{\partial t}\right)(t_i, h_j)$ and $(I_{ij})_N$ is the numerical approximation to I_{ij}, then

$$\begin{aligned} |(I_{ij})_N - I_{ij}| &= \left| \left(\frac{\partial d}{\partial h}\right)_N - \frac{\partial d}{\partial h} + \left(p\frac{\partial d}{\partial t}\right)_N - p\frac{\partial d}{\partial t} \right| \\ &\leq \epsilon\left(\left|\frac{\partial d}{\partial h}\right| + \left|p\frac{\partial d}{\partial t}\right|\right) \end{aligned}$$

Since I only consider $t \sim \sqrt{t_0^2 + 4h^2 u}$ and notice that $p = \frac{4h}{t}u$, then $|\sum I_{ij}| = o(1)$. Hence,

$$\left|\sum (I_{ij})_N - \sum I_{ij}\right| = o(1).$$

Thus 3.20 and 3.21 are the sufficient conditions to get a good approximation to the objective function for mono frequency and single reflector. Since seismic data is band-limited and layered medium has a certain number of reflectors, the numerical approximation is effective if the maximum frequency, zero-offset traveltime related to the shallowest reflector, the largest half offset, the largest half offset sampling rate, and traveltime sampling rate satisfy conditions 3.20 and 3.21.

3.4.4 Numerical Optimization

I employ the Limited Memory Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm (LBFGS) which is considered the most effective unconstrained optimization algorithm under many circumstances [Nodedal and Wright, 1999]. This algorithm is implemented with the Rice Vector Library (RVL) framework. The algorithm is coded in C/C++ with RVL-defined interfaces. The implementation sweeps through all the traces sequentially in the gather for each LBFGS iteration. The convergence of LGFGS iteration for DS-CR is reasonably fast ($O(10)$ iterations) on a 2.4 GHz Macbook using the gcc compiler suite.

Chapter 4

Results for Synthetic Data

I provide two examples illustrating the performance of the DS-CR algorithm. The data processed in the first example is a 2D synthetic seismogram shown in Chapter 4. The second example is a 2D real seismic data shown in Chapter 5. The first example compared the DS-CR with the DS-CV objective functions and also illustrated the impacts of frequency and offset increment on DS-CR for synthetic data. The second example showed the ability and deficiency of DS-CR to deal with real seismic data.

4.1 DS-CR and DS-CV

The true background velocity I used to generate this single synthetic CMP gather is constant and equal to 2000 m/s. There are eight reflectors locating at 1200m, 1600m, 2000m, 2400m, 2800m, 3200m, 3600m, and 4000m depth shown as $r(t_0)$ in Figure 4.2. Source function is synthetic Ricker wavelet with 80Hz peak frequency shown as $\omega(t)$

in Figure 4.1. The synthetic data can be computed by convolving source function $\omega(t)$ with reflectivity $r(T_0(t,h))$, which is shown in equation 2.11. The single CMP gather contains traces which are synthetic seismic signals recorded from multiple receivers displayed in Figure 4.3. 311 traces have been recorded and offset ranges from 0m to 2015m. Offset increment is 6.5m. Each vertical trace represents one receiver's response. Every trace is a timing line at 2ms intervals and has 1301 timing samples. In this example, I compared the DS-CR and DS-CV objective functions described in the preceding section, using this synthetic CMP gather and different velocities.

To illustrate that my objective function is a smooth function with respect to background velocity and has less noise, I use velocity displayed in Figure 4.2 as the reference velocity v_0 and constant true velocity as the target velocity v_1. Velocity v_k is the convex combination of v_0 and v

$$v_k = v_0 + k \times (v_1 - v_0),$$

where k ranges from 0.94 to 1.1 and Δk is 0.005. Left figure in 4.4 is the DS-CR objective function $J_{DS-CR}[v]$ with respect to convex combination coefficient k. It is a smooth curve with only one global minimizer where $k = 1$, i.e. $v_k = v_1$. Right figure in 4.4 shows that there exists local minimizers in the DS-CV approach. The objective function of DS-CV has more noise than my approach. As explained in the previous chapter, the reason for this error to occur is that the DS-CV approach inevitably involves interpolation of oscillatory seismic data, while the numerical approximation to the objective function for DS-CR only involves interpolation of smooth function.

It can be noticed in the two figures that the DS-CR objective function values are always larger than DS-CV. This result is caused by discretization error, which will be explained in detail in the next chapter.

4.2 Impacts of Frequency and Offset Increment on the DS-CR Accuracy

This example shows that what impacts of frequency and offset increment have on the DS-CR accuracy. The data is generated with the same constant velocity model (2000 m/s) and eight reflectors displayed in Figure 4.2. For different frequencies and offset increments, I compared the data flatness after NMO correction using velocity estimated by DS-CR. Since $\Delta t = 2ms$ and maximum frequency is 40Hz, then

$$\max_{\Delta t, f} \left(\frac{\Delta t^2}{2} \pi^2 f^2 \right) = 0.0316.$$

If I choose any $\epsilon \geq 0.0316$, then first condition 3.20 will be satisfied. I choose $\epsilon = 0.05$. Since maximum half offset $h_{max} = 1000m$, minimum zero-offset traveltime $t_{0min} = 0.6s$, maximum frequency $f_{max} = 40Hz$, RMS square slowness $u = 2.5 \times 10^{-7} s^2/m^2$ and error bound $\epsilon = 0.05$, then I have error estimated by 3.21

$$\begin{aligned} &\max_{h, \Delta h, f, t_0} \left(\frac{\Delta h^2}{6} \left(\frac{16\pi^2 h^2 u^2 f^2}{T^2} + \frac{6\pi t_0^2 u f}{T^3} + \frac{3 t_0^2 u}{T^4} \right) \right) \\ &= \left(800\pi^2 + 30\sqrt{2}\pi + 0.75 \right) \frac{\Delta h^2}{24} \leq \epsilon. \end{aligned}$$

Sufficient condition for discretization error smaller than 0.05 is $\Delta h \leq 5.1m$, i.e. if offset sampling interval is less than $10.2m$, I should be able to get relatively good NMO correction. In this numerical example, I use eight different offset sampling intervals which are 40m, 30m, 20m, 15m, 13m, 12m, 11m, and 10m. The peak frequency of synthetic source Ricker wavelet is 20Hz. The synthetic CMP gathers are shown in Figure 4.5 and 4.6. I computed the final velocities using DS-CR and apply NMO correction to the original seismic data with those final velocities. From NMO corrected data shown in Figure 4.7 and 4.8, it is required that offset interval Δx has to be smaller than 11m in order to get a flat NMO correction which is consistent with the sufficient condition. For this experiment, the data can be flattened if and only if the offset sampling interval is smaller than 11m.

Since maximum half offset $h_{max} = 1000m$, minimum zero-offset traveltime $t_{0min} = 0.6s$, half offset sampling interval $\Delta h = 20m$, RMS square slowness $u = 2.5 \times 10^{-7} s^2/m^2$ and error bound $\epsilon = 0.05$, then condition 3.21 is

$$\max_{h,\Delta h,f,t_0} \left(\frac{\Delta h^2}{6} \left(\frac{16\pi^2 h^2 u^2 f^2}{B^2} + \frac{6\pi t_0^2 u f}{B^3} + \frac{3t_0^2 u}{B^4} \right) \right)$$
$$= \left(2\pi^2 f_{max}^2 + 1.5\sqrt{2}\pi f_{max} + 0.75 \right) \frac{0.02^2}{24} \leq \epsilon.$$

Second sufficient condition for discretization error smaller than 0.05 is $f_{max} \leq 10.1$. This maximum frequency f_{max} also satisfies the first sufficient condition 3.20. So if peak frequency fp is less than $5.05Hz$, I should be able to get relatively good NMO correction. In this numerical example, I use eight different peak frequencies which are 25Hz, 20Hz, 15Hz, 10Hz, 8Hz, 7Hz, 6Hz, and 5Hz. The offset sampling rate is

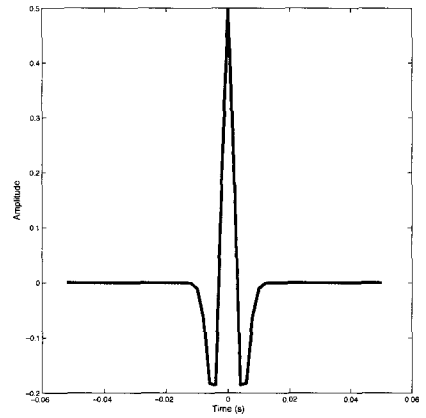

Figure 4.1: Source time function $w(t)$: Ricker wavelet with peak frequency 80Hz

40m. The synthetic CMP gathers are shown in Figure 4.9 and 4.10. I computed the final velocities using DS-CR and apply NMO correction to the original seismic data with those final velocities. From NMO corrected data shown in 4.11 and 4.12, it is required that fp to be smaller than 6Hz in order to get a flat NMO correction which is consistent with the sufficient condition. In this experiment, the data can be flattened if and only if the peak frequency is smaller than 6Hz.

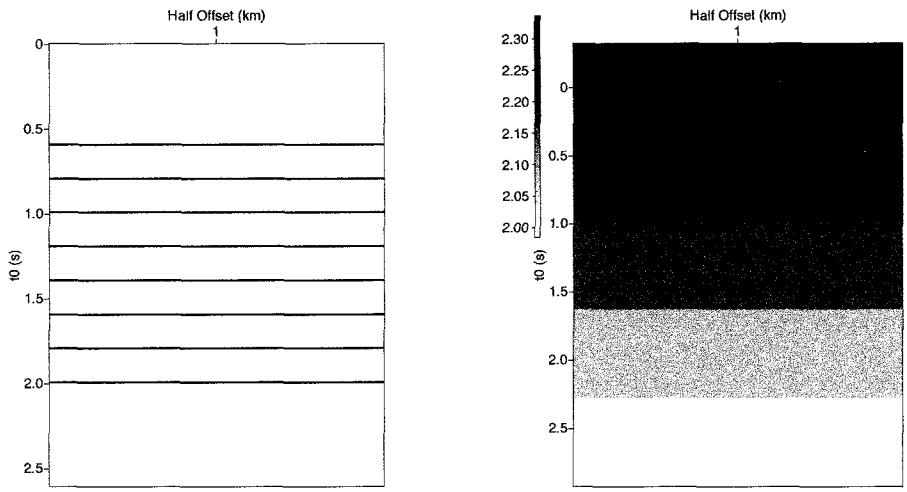

Figure 4.2: Left: Relative velocity perturbation (reflectivity $r(t_0)$). Right: Reference velocity $v(t_0)$.

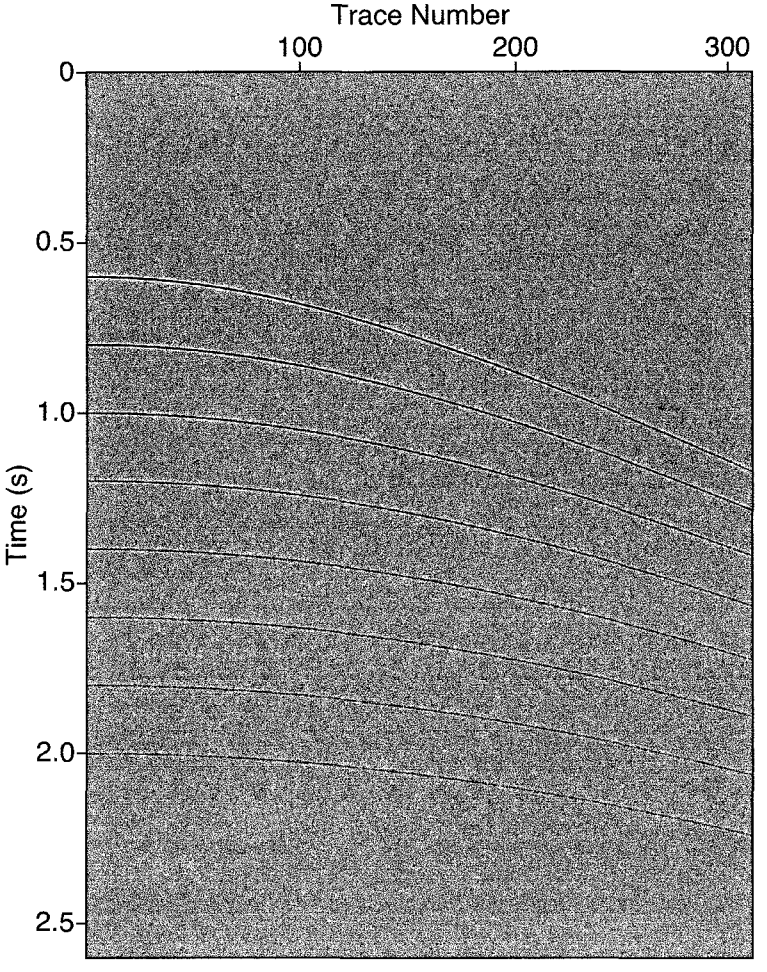

Figure 4.3: Synthetic seismic data for a single CMP gather. There are 311 traces and offset sampling rate is 6.5 m. There are 1301 timing samples and time sampling rate is 2 ms.

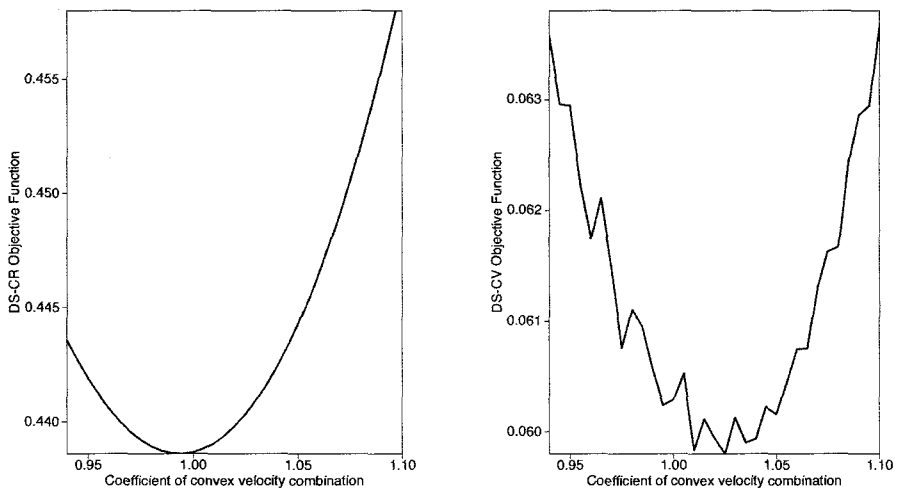

Figure 4.4: Left: the DS-CR objective function with respect to the convex velocity combination parameter k. Right: the DS-CV objective function with respect to the convex velocity combination parameter k.

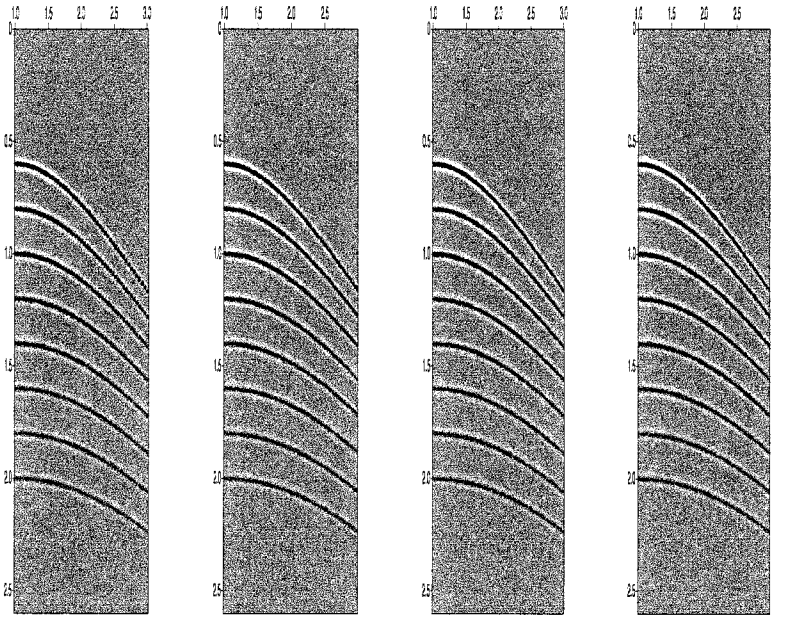

Figure 4.5: Different synthetic CMP gathers generated using the same source wavelet with peak frequency $fp = 20Hz$, and different offset sampling rate Δx. For each gather, there are 1301 timing samples and time sampling rate is 2 ms. Left:$\Delta x = 40m$; Middle left:$\Delta x = 30m$; Middle right:$\Delta x = 20m$; Right:$\Delta x = 15m$

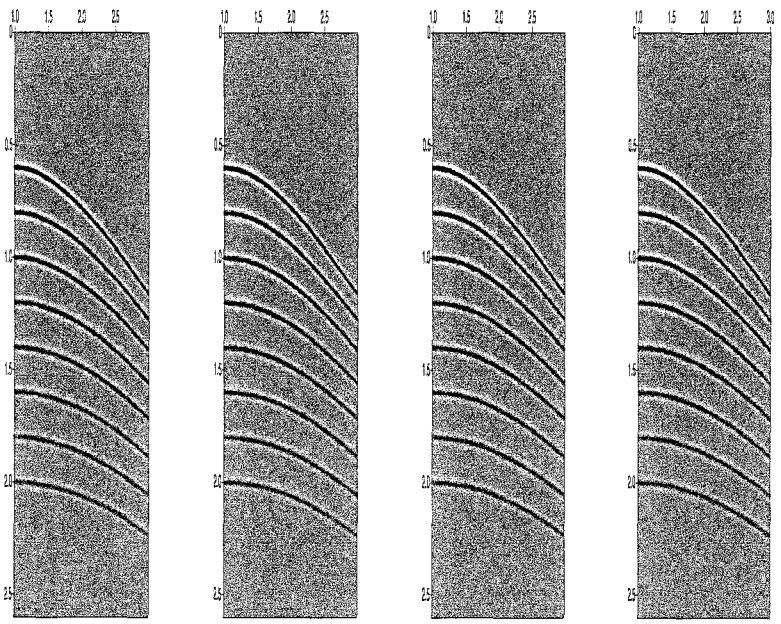

Figure 4.6: Different synthetic CMP gathers generated using the same source wavelet with peak frequency $fp = 20Hz$, and different offset sampling rate Δx. For each gather, there are 1301 timing samples and time sampling rate is 2 ms. Left:$\Delta x = 13m$; Middle left:$\Delta x = 12m$; Middle right:$\Delta x = 11m$; Right:$\Delta x = 10m$

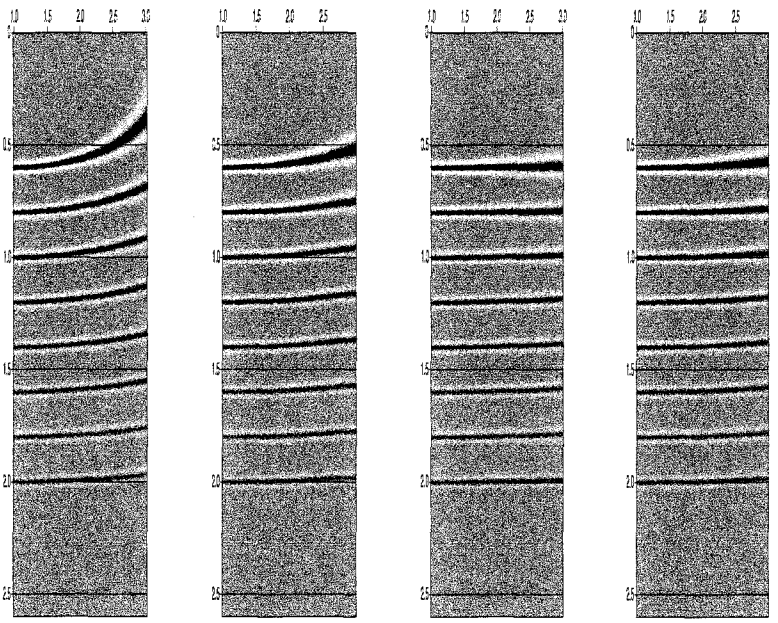

Figure 4.7: NMO corrections for different synthetic CMP gathers (Figure 4.5) generated using the same source wavelet with peak frequency $fp = 20Hz$, and different offset sampling rate Δx. Velocities are computed by DS-CR. Left:$\Delta x = 40m$; Middle left:$\Delta x = 30m$; Middle right:$\Delta x = 20m$; Right:$\Delta x = 15m$

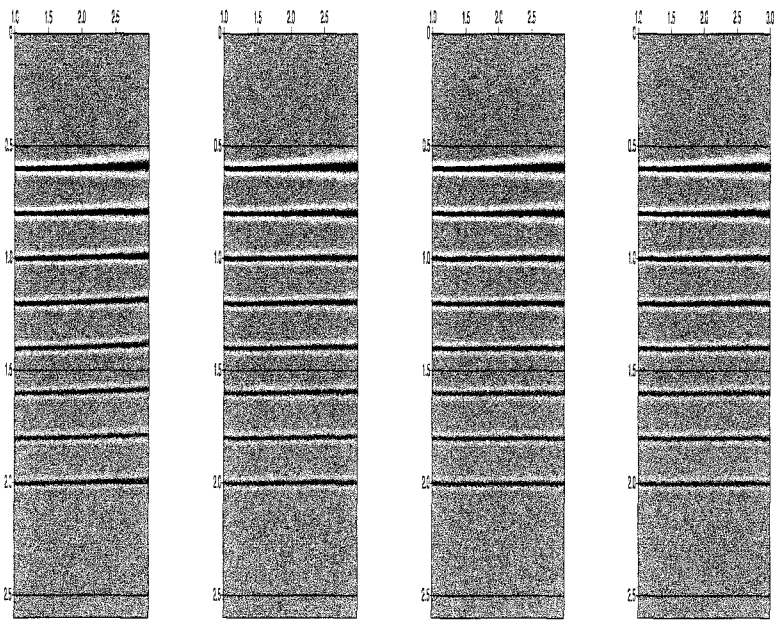

Figure 4.8: NMO corrections for different synthetic CMP gathers (Figure 4.6) generated using the same source wavelet with peak frequency $fp = 20Hz$, and different offset sampling rate Δx. Velocities are computed by DS-CR. Left:$\Delta x = 13m$; Middle left:$\Delta x = 12m$; Middle right:$\Delta x = 11m$; Right:$\Delta x = 10m$

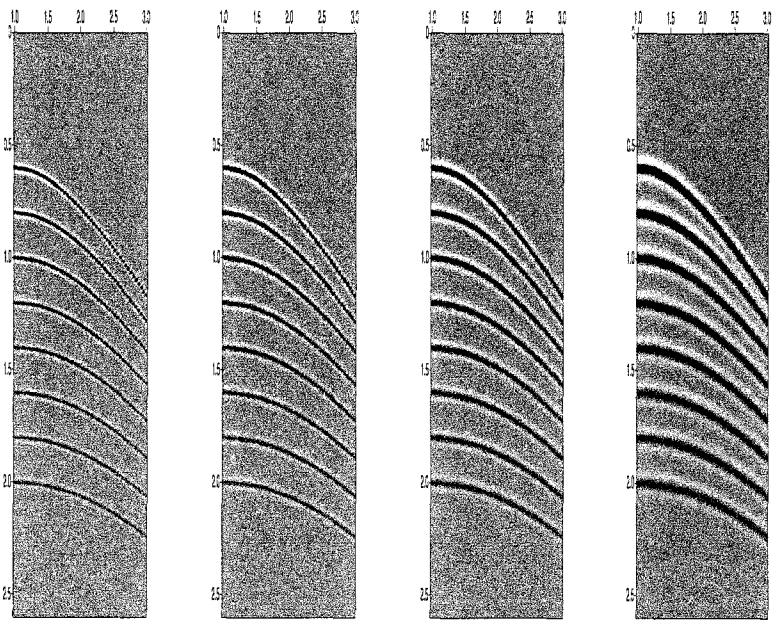

Figure 4.9: Different synthetic CMP gathers generated using source wavelet with different peak frequency fp, and the same offset sampling rate $\Delta x = 40m$. For each gather, there are 1301 timing samples and time sampling rate is 2 ms. Left:$fp = 25Hz$; Middle left:$fp = 20Hz$; Middle right:$fp = 15Hz$; Right:$fp = 10Hz$

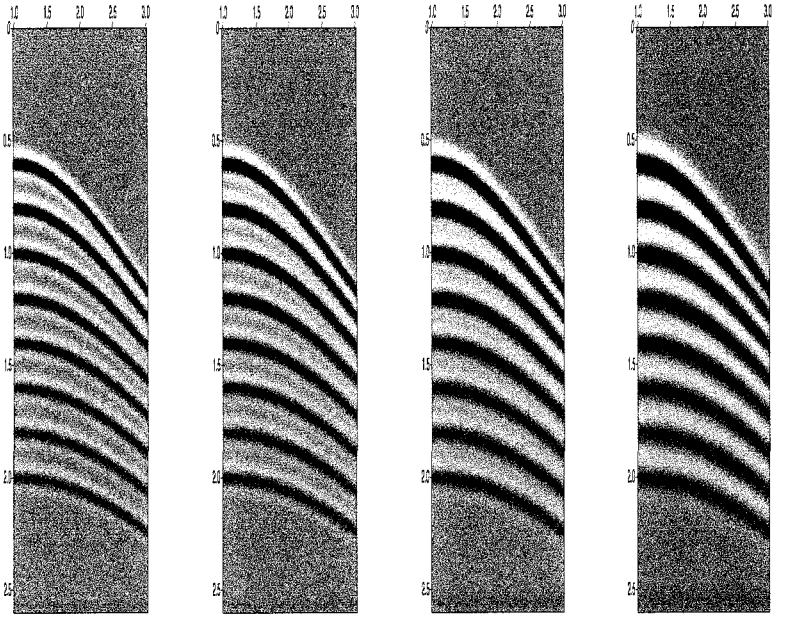

Figure 4.10: Different synthetic CMP gathers generated using source wavelet with different peak frequency fp, and the same offset sampling rate $\Delta x = 40m$. For each gather, there are 1301 timing samples and time sampling rate is 2 ms. Left:$fp = 8Hz$; Middle left:$fp = 7Hz$; Middle right:$fp = 6Hz$; Right:$fp = 5Hz$

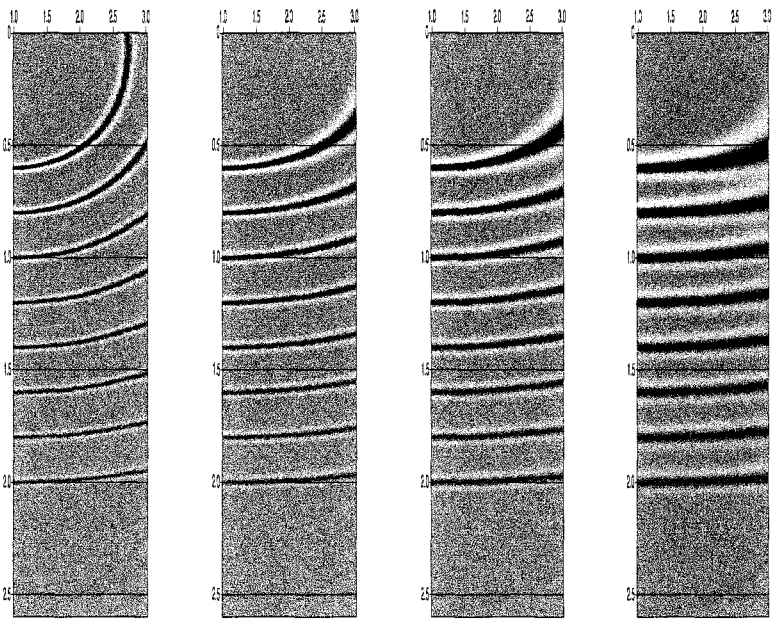

Figure 4.11: NMO corrections for different synthetic CMP gathers (Figure 4.9) generated using source wavelet with different peak frequency fp, and the same offset sampling rate $\Delta x = 40m$. Left:$fp = 25Hz$; Middle left:$fp = 20Hz$; Middle right:$fp = 15Hz$; Right:$fp = 10Hz$

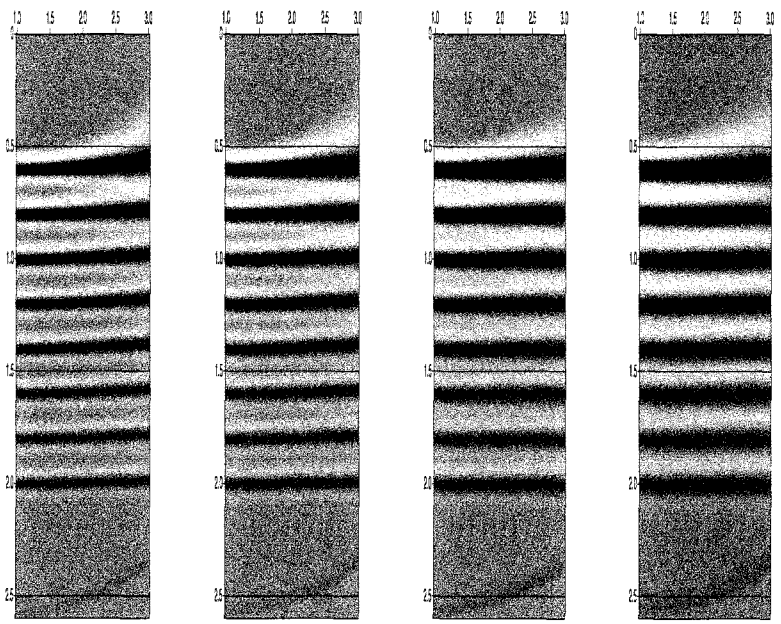

Figure 4.12: NMO corrections for different synthetic CMP gathers (Figure 4.10) generated using source wavelet with different peak frequency fp, and the same offset sampling rate $\Delta x = 40m$. Left:$fp = 8Hz$; Middle left:$fp = 7Hz$; Middle right:$fp = 6Hz$; Right:$fp = 5Hz$

Chapter 5

Results for Real Data

The processed data in this example is 2D marine data recorded from the North Sea released by Shell. This example presents the ability and deficiency of DS-CR to deal with real seismic data, compared to CS and DS-CV. In this example, I use the CMP gather displayed in Figure 5.1 as input data. Figure 5.1 is a single CMP gather containing wiggle-line traces which are seismic signals recorded from multiple hydrophones. Traces (vertical lines) are timing lines at 2 ms intervals. Each trace represents one hydrophone's response and has 1001 samples. 50 traces have been recorded and offset ranges from 185 m to 1965 m.

5.1 Stability

As explained in section 3.4.1, I employ piecewise linear interpolation of the nodal values for velocity representation in order to produce relatively smooth velocity model.

I define the fixed velocity nodes at 5 depths raging from 0 to 2 s, at intervals of 0.5 s, because 5 nodes representation is sufficient for this example to estimate relatively smooth final velocity that satisfies the optimization criterion. I compare the sensitivity of final velocity estimates to different initial velocity models between DS-CR and DS-CV first. Given different initial velocities, the DS-CR final estimates are very close, particularly between the surface and $t_0 = 1.5$ s. Figure 5.2 displays this result. Figure 5.3 shows the estimated velocities by applying DS-CV for the same input data with four different initial models. Those final velocities have obvious variations. Figure 5.2 and 5.3 illustrated the tendency of DS-CR to produce equivalent velocities independent of initial velocity guess so that DS-CR tends to be more stable.

5.2 Flatness

Though the final velocities computed by DS-CR and DS-CV are different, the NMO-corrected data are similarly flattened by those final velocities, shown in Figure 5.4 and 5.5. With careful comparison, it can be noticed that Figure 5.5 seems to be better corrected. The reason is that DS-CR involves discretization error in the approximation to its objective function in this example. The DS-CR objective function value is almost 3 times larger than the DS-CV objective function value for the same velocity model. As explained in the previous chapter, the key idea of normal moveout based differential semblance is to minimize the l^2-norm of the DS-CR moveout derivative $\left(\frac{\partial d}{\partial h} + p\frac{\partial d}{\partial t}\right)(t,h)$ or the DS-CV moveout derivative $\frac{\partial}{\partial h}d(T(t_0,h),h)$. I apply a change

of variables map $T_0(t,h)$ for the DS-CV moveout derivatives to transfer them from t_0-h domain to $t-h$ domain so that consistent with the DS-CR moveout derivatives. In order to compare the numerical error, I illustrates the moveout derivatives using two approach in Figure 5.6 and 5.7 with the same velocity (blue final estimate in Figure 5.3). Both of the DS-CR and the DS-CV moveout derivatives should be close to zero since they are exactly the same in the continuous form and the velocity model used to compute the moveout derivatives is assumed to be a good estimate. But DS-CR exhibits more numerical noise in the computation of moveout derivatives. They differ mostly in the shallow and large offset panels in the domain. This phenomenon is consistent with the sufficient condition 3.21 which describes that smaller t_0 and larger h would increase the error bound value. After observing the frequency spectrum shown in Figure 5.8, we can find out that the peak frequency is about $33Hz$. The offset sampling rate which is around $38m$ is deficient according to condition 3.21 since it can only guarantee that the discretization error to the half-offset derivative is smaller than 150%. This discretization error pollutes the DS-CR moveout derivatives and the DS-CR objective function as well. As a consequence of wrong objective function estimation, the computed gradient cannot provide the descending direction and causes premature termination during iterative optimization. In the end of the optimization log, it records that the line search failed, steepest descent restarted, and no smaller objective function value could be found. DS-CR is not able to find the exact global (local) minimum in this example. Therefore, DS-CR is not able to com-

pletely flatten the data with its final velocity estimate in this example and shallow panels are less flattened than deep panels. I may conclude that the reason for which Figure 5.4 seems to be incompletely flattened after NMO correction is the insufficient offset sampling issue, explained in the previous synthetic example in detail. I might overcome this deficient sampling issue by applying a lowpass filter in order to decrease the frequencies. Satisfactory frequency should be less than 10Hz according to condition 3.21. But this idea cannot be applied for this data set since there is no data for frequency under $10Hz$ shown in Figure 5.8.

5.3 Global Minimum

In order to compare DS and CS, I use thin black initial velocity displayed in Figure 5.2 as the reference velocity v_0 and the thick black final velocity as the target velocity v_1. Velocity v_k is the linear combination of v_0 and v_1

$$v_k = v_0 + k \times (v_1 - v_0),$$

where k ranges from 0 to 2. The increment Δk is 0.002. Left figure in 5.9 is the DS-CR objective function $J_{DS-CR}[v_k]$ with respect to k. It is a convex and smooth curve with only one global minimizer where $k = 0.94$. Right Figure in 5.9 suggests that there is no local minimizer no matter how much the Left figure in 5.9 has been magnified. The final velocity estimated by DS-CR is v_1, while the global minimizer is $v_{0.94}$. The optimization fails to find the exact global minimum, but ends up with some

nearby velocity. This again explains why the data cannot be flattened completely in Figure 5.4 due to the premature termination. Compared to DS-CR, CS does not exhibit this great property in terms of global convexity (concavity) of the objective function, which has been mentioned in section 3.1. Figure 5.10 illustrates that classical semblance has local maximizers.

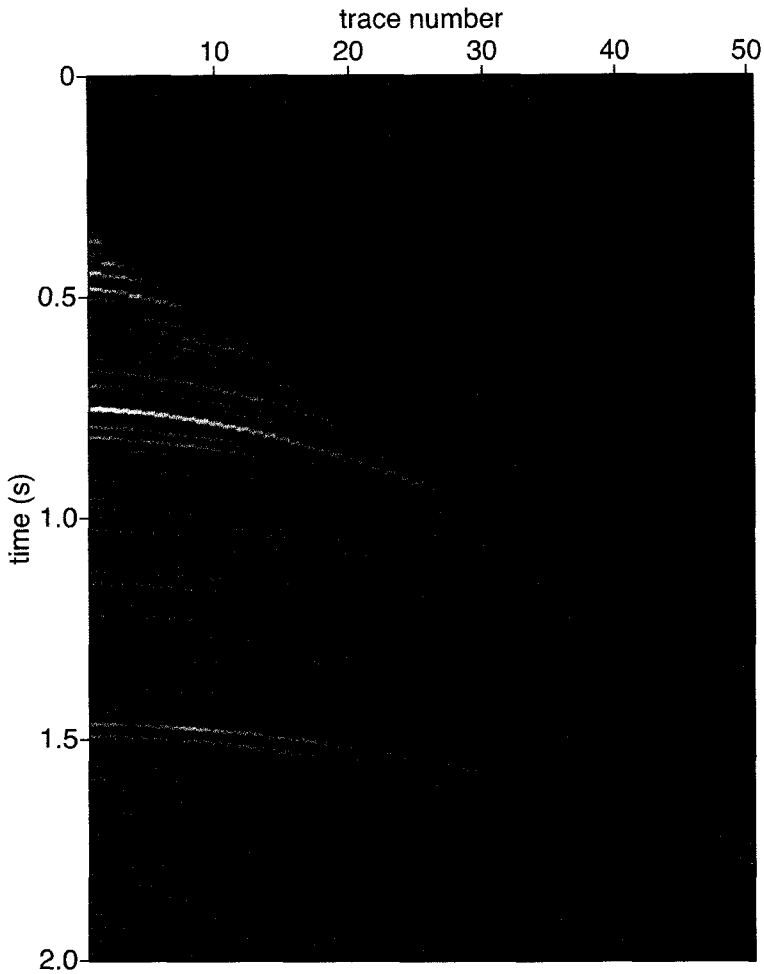

Figure 5.1: CMP number -11 from the 2D marine data released by Shell. There are 50 traces and offset sampling rate around 38 m. There are 1001 timing samples and time sampling rate is 2 ms.

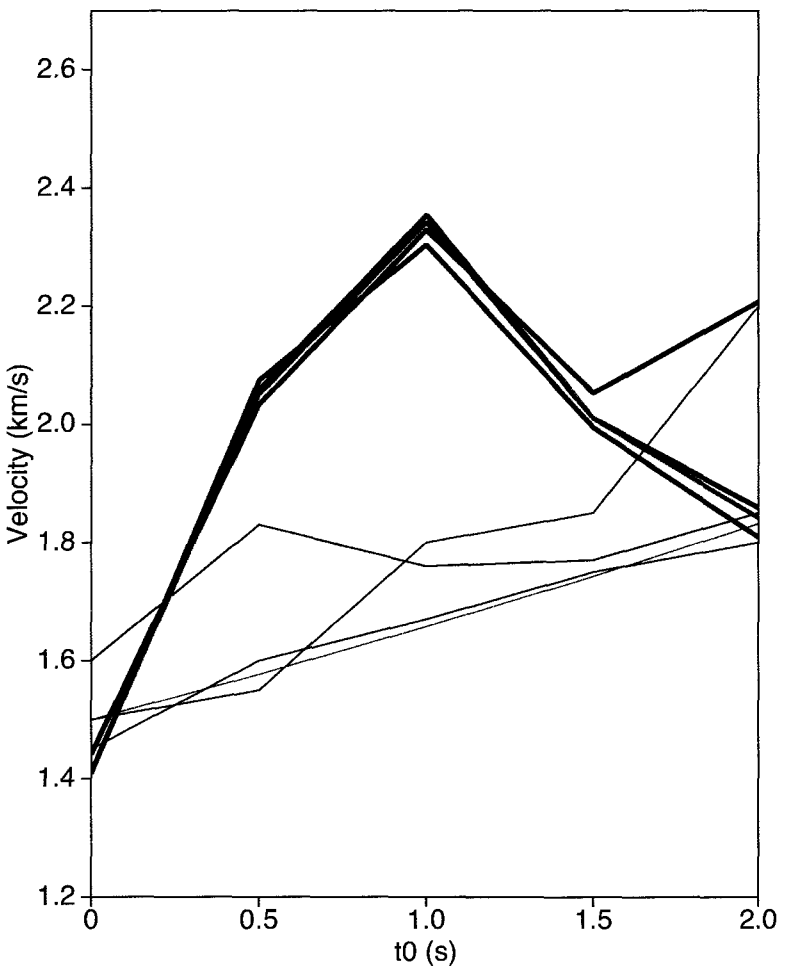

Figure 5.2: Instability of DS-CR. Narrow solid lines: initial velocity estimates. Thick solid lines: corresponding final velocity estimates. Input data: CMP -11

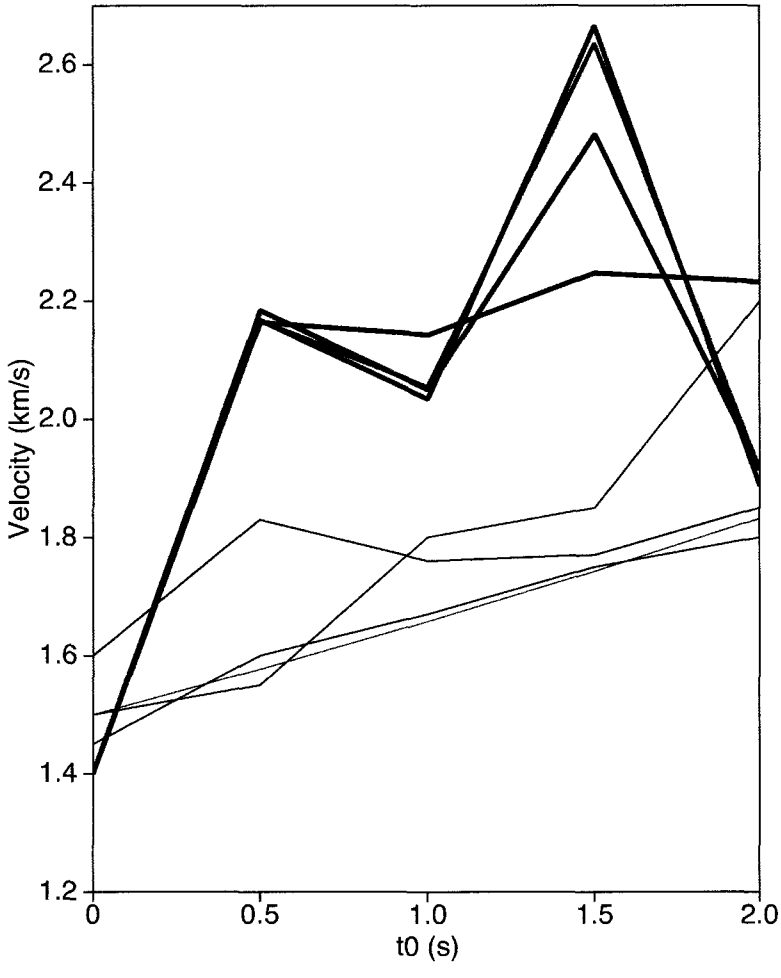

Figure 5.3: Instability of DS-CV. Narrow solid lines: initial velocity estimates. Thick solid lines: corresponding final velocity estimates. Input data: CMP -11

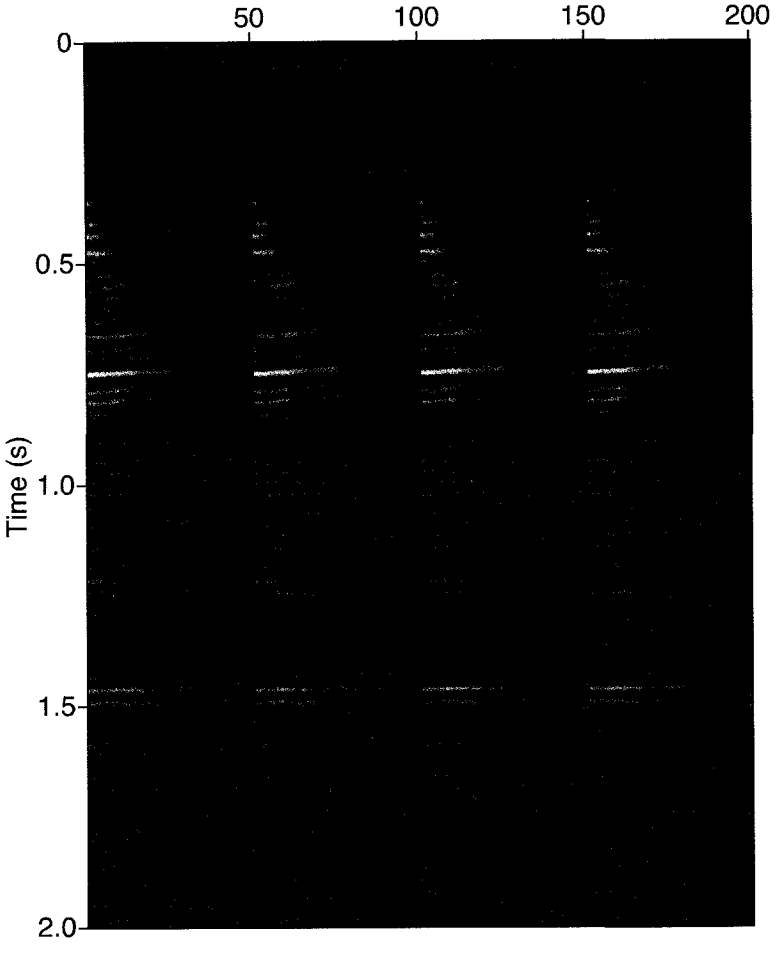

Figure 5.4: CMP -11 after normal moveout correction using the final velocity estimates shown in Figure 5.2

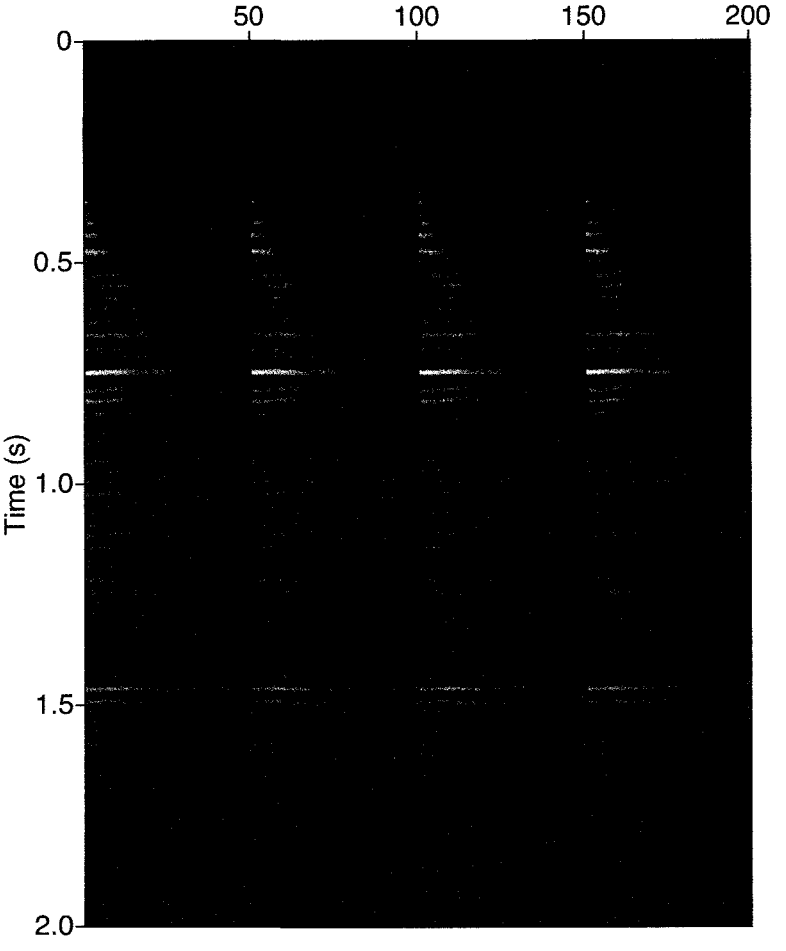

Figure 5.5: CMP -11 after normal moveout correction using the final velocity estimates shown in Figure 5.3

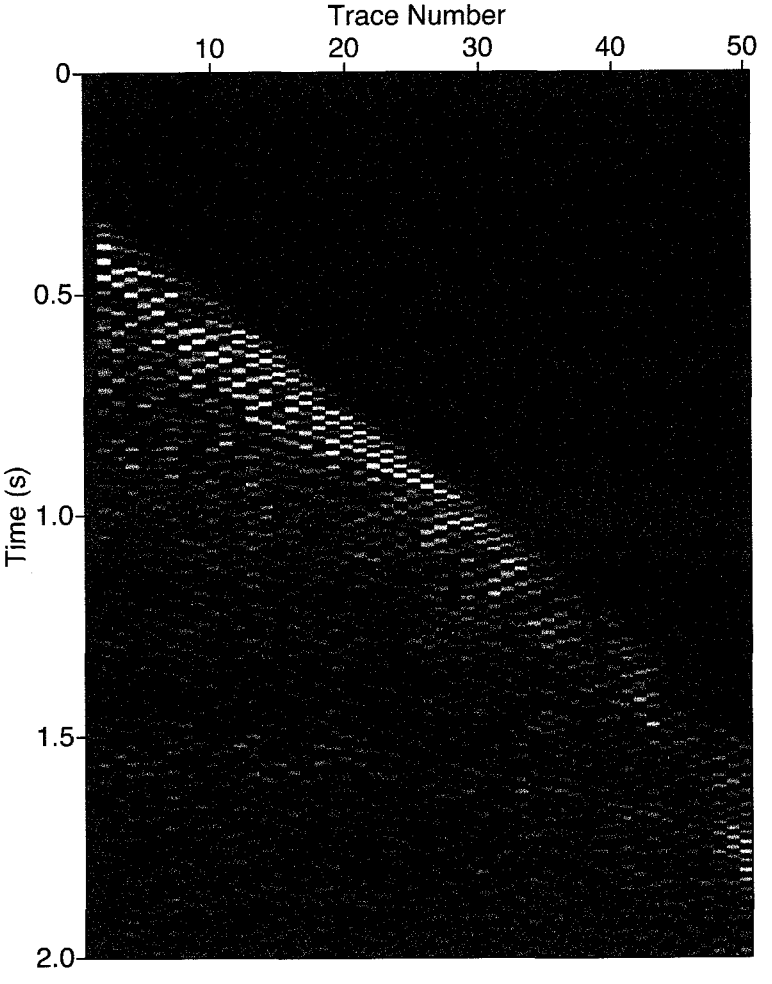

Figure 5.6: The DS-CR moveout derivatives using the thick blue velocity in Figure 5.3 $\left(\frac{\partial d}{\partial h} + p\frac{\partial d}{\partial t}\right)(t, h)$

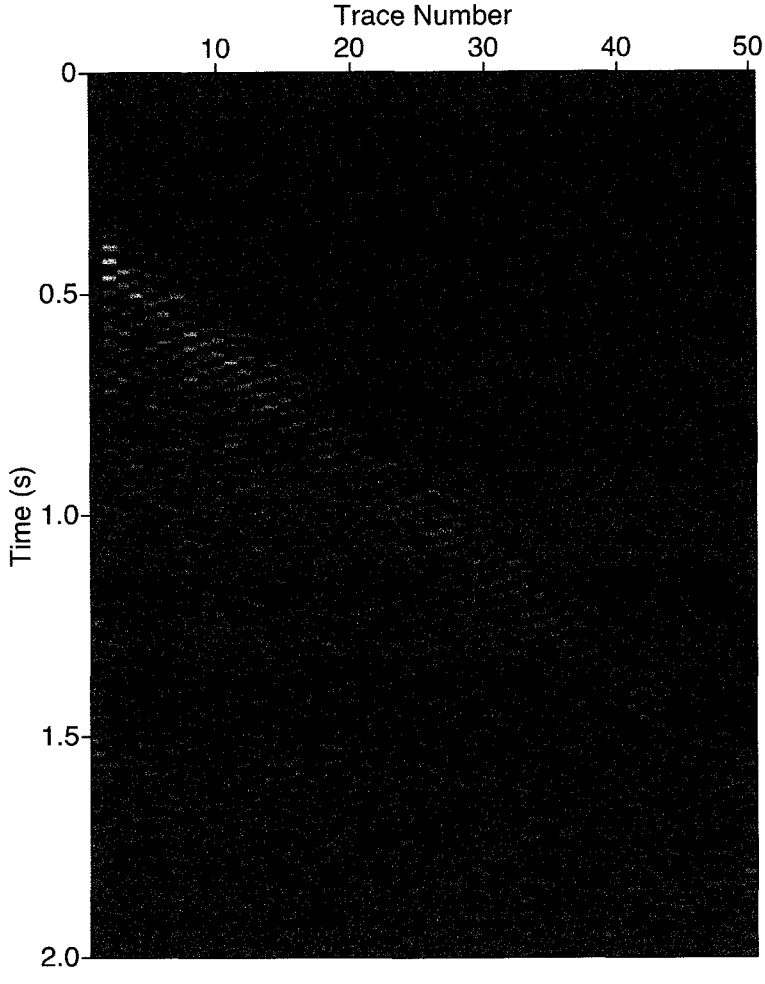

Figure 5.7: The DS-CV moveout derivatives using the thick blue velocity in Figure 5.3 $\frac{\partial}{\partial h} d(T(t_0, h), h)|_{t_0 = T_0(t,h)}$

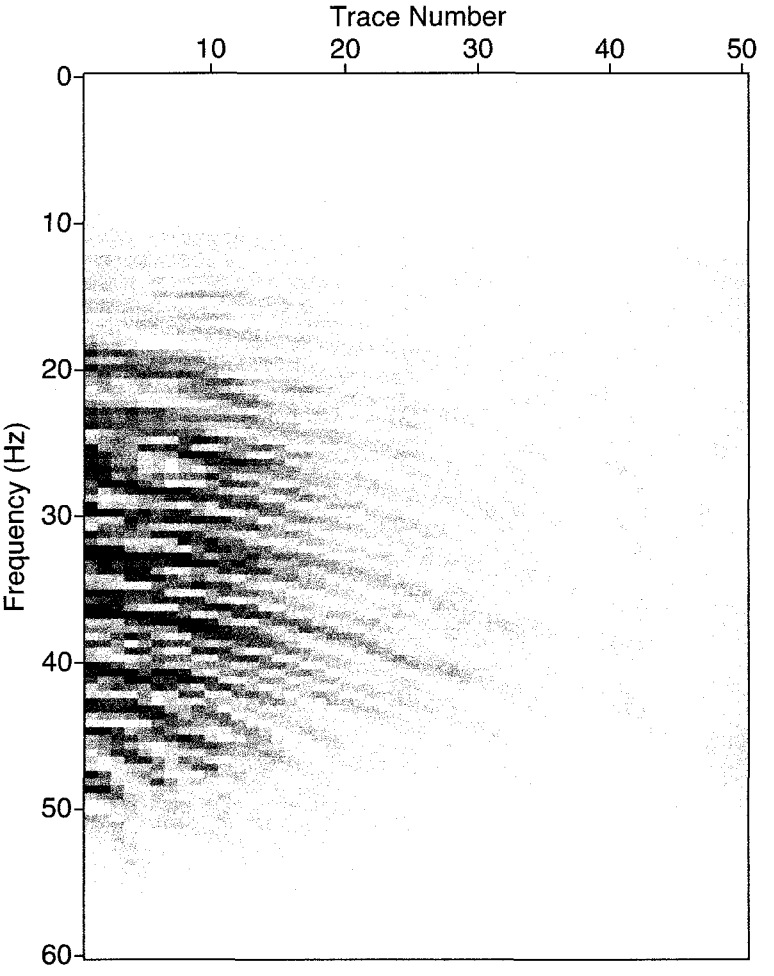

Figure 5.8: Frequency spectrum for CMP number -11 from the 2D marine data released by Shell

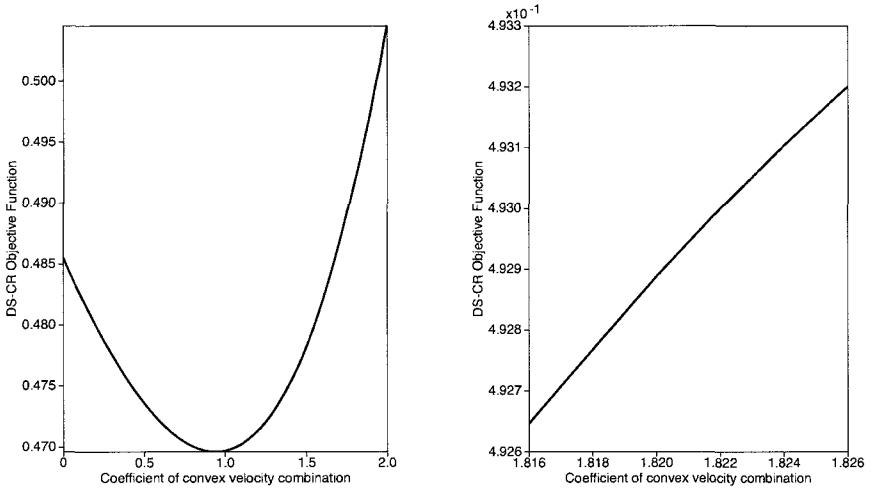

Figure 5.9: Left: the DS-CR objective function $J_{DS-CR}[v_0 + k(v_1 - v_0)]$ with respect to the convex velocity combination parameter k. The parameter k ranges from 0 to 2. The increment Δk is 0.002. Reference velocity v_0 is the thin black velocity model in Figure 5.2. Target velocity v_1 is the DS-CR final velocity estimate with v_0 as the initial guess (thick black velocity model in 5.2). Right: the zoomed DS-CR objective function $J_{DS-CR}[v_0 + k(v_1 - v_0)]$ with respect to the convex combination parameter k. The parameter k ranges from 1.55 to 2.

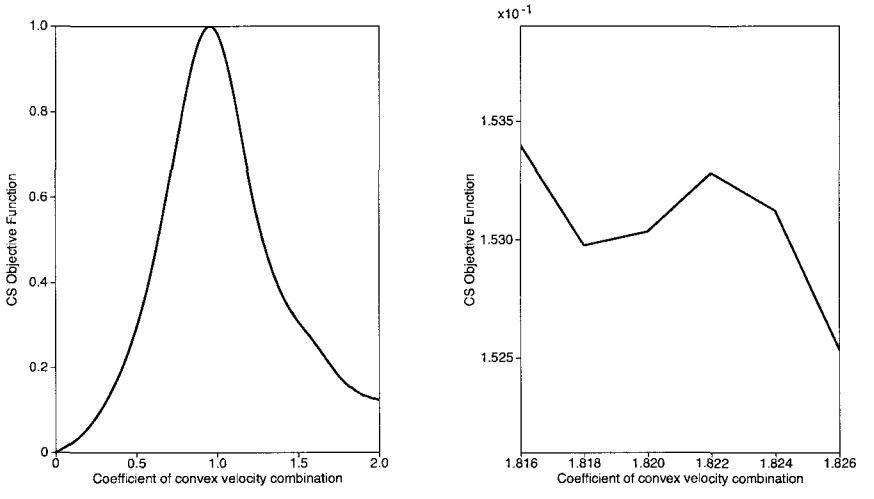

Figure 5.10: Left: CS objective function $J_{CS}[v_0 + k(v_1 - v_0)]$ with respect to the convex velocity combination parameter k. The parameter k ranges from 0 to 2. The increment Δk is 0.002. Reference velocity v_0 is the thin black velocity model in Figure 5.2. Target velocity v_1 is CS final velocity estimate with v_0 as the initial guess. Right: zoomed CS objective function $J_{CS}[v_0 + k(v_1 - v_0)]$ with respect to the convex combination parameter k. The parameter k ranges from 1.816 to 1.826.

Chapter 6

Conclusions and Discussions

6.1 Conclusions

I have described a new NMO-based implementation of DSVA and shown 2D examples of its use. The smoothness and convexity of the DS-CR objective function has been shown in the previous chapter, so my approach allows automatic velocity update by means of local search methods. When the sufficient conditions 3.20 and 3.21 are satisfied, this implementation has the capability to estimate 1D, 2D, or 3D velocity models within the limits that the medium is mild lateral heterogeneous and the data are dominated by primary events, which are the same model assumptions applied to DS-CV. DS-CR and DS-CV have similar capacity of flattening the seismic data. DS-CR is able to converge to a reasonable final velocity in a few iterations (O(10)), similar convergence rate as DS-CV.

The previous approach DS-CV [Li and Symes, 2007] involved the interpolation of the oscillatory seismic data and as a result the interpolation error prevented the gradient from descending after several iterations and led to some large variance in the estimated interval velocities, which are local minimizers. This thesis develops an alternative approach based on Li's thesis. My approach employs a new discretization method to approximate the same continuous DS objective function as DS-CV and a new computation of the gradient. By applying chain rule, this alternative computation of the objective function interpolates only smooth functions instead of interpolating oscillatory data and the gradient has less interpolation error as well, so that this approach has less numerical interpolation error than the previous implementation and the DS-CR objective function is always convex and smooth. DS-CR is more stable than DS-CV since it is less sensitive to the initial velocity variations. However, if the frequency is too high or the offset sampling rate is too big, DS-CR will have larger discretization error to approximate the h partial derivative and lead to premature termination during optimization. In order to overcome this defect, one way is to increase the offset samples during acquisition of data. The other way is to apply a lowpass filter before velocity analysis procedure. Although both approaches are similarly good at flattening the hyperbola of primary events for mild laterally heterogeneous medium, the insensitivity of DS-CR to a prior knowledge about background velocity is its greatest strength compared to DS-CV when the sufficient conditions are satisfied.

6.2 Numerical Issue

The numerical issue revealed by this new implementation is the sampling issue. The sampling rate is an important factor, because if the time sampling intervals are too far apart, the true form of the signals is not preserved and the high frequencies are lost. In practice, modern equipments have the capacity to record signals at very small time sampling rates. Multiple hydrophones are arranged in an array in order to generate different traces. In real seismic acquisition, the half offset h is often undersampled due to the high cost of equipments. Notice that the distance between neighboring traces is about 38 m, while the time sample rate is only 2 ms. The use of the cell centered difference to approximate the h partial derivative is a likely source of relatively large discretization error due to deficient samples of offset. Such error may prevent the gradient from finding the descending direction after several iterations and lead to premature termination of iterative optimization. As the synthetic example shown in the preceding chapter, the image after NMO correction using the DS-CR approach was not flat any more when I increased the offset sampling rate. This issue can be overcome by applying a lowpass filter to decrease the frequency as shown in the synthetic example. Thus in order to make DS-CR effective, either small offset sampling rate or low frequency data is required for the real seismic data.

A better algorithm should be employed here to compute the offset derivative. I propose an alternative option in numerical implementation which might offer better performance. I split the single optimization problem into several sub-problems. Each

sub-problem is an optimization problem within certain velocity envelop. Use the initial velocity estimate v_0 of each velocity envelop and the DS-CV method to compute moveout derivative J_0. Then estimate

$$\frac{\partial d}{\partial h} = J_0 - p_0[v_0]\frac{\partial d}{\partial t}.$$

More accurate velocity can be estimated by updating the velocity within each envelop using DS-CR. Once I get the optimal velocity for current envelop, $\frac{\partial d}{\partial h}$ can be approximated more accurately using the current optimal velocity as the initial velocity guess, which is the beginning of the next envelop. This new idea seems can make some improvement to overcome the undersampled problem occurred in DS-CR.

6.3 Coherent Noise

Another issue to be considered is the coherent noise. The consequence of inadequate multiple suppression is shown in the example presented in the previous chapter, which is a shortcoming of partial linearization. The presence of coherent noise degrades the quality of velocity estimates. DSVA overcorrects the primaries and undercorrects the multiples. One way to get rid of the coherent noise is to apply some demultiple filters before DSVA. Another way is to develop another approach to velocity estimation taking into account the nonlinear scattering. See [Symes, 2008] for an overview of recent research on nonlinear waveform inversion and differential semblance.

6.4 Layered Model Constraint

In reality, the media usually has complex subsurface structure with sharp lateral velocity variations. Due to the shortcomings of the convolution model and the hyperbolic moveout approximation, the DSVA via NMO correction cannot handle multipathing and complex geological structure. DSVA via Reverse Time Migration, should provide the most accurate and reliable velocity estimation without dip limitation.

APPENDIX A: Classical Semblance's Equivalency to Least Squares

Turn the linearized inverse problem into a least squares problem: given CMP data $d(t,h)$, find $v(t_0)$, $r(t_0)$ so that

$$\min_{v,r} \; J_{OL}[v,r](t,h) = \|F[v]r(t,h) - d(t,h)\|^2 .$$

Assume the point source time function $\omega(t) = \delta(t)$ shown in chapter 2, then

$$F[v]r(t,h) = r(T_0(t,h)),$$

by equation 2.3.

$$\begin{aligned}
\|r(T_0(t,h)) - d(t,h)\|^2 &= \int\int dt\ dh\ (r(T_0(t,h)) - d(t,h))^2 \\
&= \|d\|^2 + \int\int dt_0\ dh\ \frac{\partial T}{\partial t_0}(t_0,h) \times (r(t_0)^2 - 2r(t_0)d(T(t_0,h),h)) \\
&= \|d\|^2 + \int dt_0\ w(t_0)r(t_0)^2 - 2\int dt_0\ r(t_0)\int dh\ \frac{\partial T}{\partial t_0}(t_0,h)d(T(t_0,h),h)
\end{aligned}$$

Then

$$J[v,r] = \| \ d \ \|^2 + < wr, r > - \ 2 < r, S >,$$

where S is the weighted stacking power

$$S[v](t_0) = \int dh\ \frac{\partial T}{\partial t_0}(t_0,h)d(T(t_0,h),h),$$

and

$$w[v](t_0) = \int dh\ \frac{\partial T}{\partial t_0}(t_0,h).$$

Since S, w only depend on v, then if v is fixed, we can get the optimal $r = \frac{1}{w}S$

$$\min_{v,r} \; J_{OL}[v,r] = \; \|d\|^2 - < \frac{1}{w}S, S >$$

$$\Longleftrightarrow \max_{v} \; J_{CS}[v] = < \frac{1}{w}S, S >$$

Then the classical semblance turns out to be equivalent to the least squares.

APPENDIX B: Derivation of Analytic Expression for the DS Gradient

In this appendix, I compute the analytic expression for the gradient. For a single CMP, the objective function is

$$J[v] = \frac{1}{2}\int dh \int dt \left| \left(p\frac{\partial d}{\partial t} + \frac{\partial d}{\partial h} \right)(t,h) \right|^2, \tag{B-1}$$

where the slowness

$$p(t,h) = \frac{4h}{t} u(T_0(t,h)). \tag{B-2}$$

Denote the directional derivative of v-dependent quantities in the direction δv by the prefix δ. Denote

$$I(t,h) = \left(p\frac{\partial d}{\partial t} + \frac{\partial d}{\partial h} \right)(t,h).$$

Then

$$J = \frac{1}{2}\langle I, I \rangle .$$

$$\begin{aligned} \delta J &= \langle \delta I, I \rangle \\ &= \langle DI\delta v, I \rangle \\ &= \langle \delta v, (DI)^* I \rangle \end{aligned}$$

Thus $\nabla_v J = (DI)^* I$.

From equation B-1,

$$\delta J[v] = \int dh \int dt\ \delta p(t,h) \frac{\partial d}{\partial t}(t,h) \left(p\frac{\partial d}{\partial t} + \frac{\partial d}{\partial h} \right)(t,h). \tag{B-3}$$

From equation B-2, it follows that

$$\delta p(t,h) = \frac{4h}{t}\left(\delta u(T_0(t,h)) + \frac{\partial u}{\partial t_0}(T_0(t,h))\delta T_0(t,h)\right). \tag{B-4}$$

Using the relation $T(T_0(t,h),h) = t$ and the definition of hyperbolic two-way travel time,

$$\delta T_0(t,h) = -\frac{\delta T(T_0(t,h),h)}{\frac{\partial T}{\partial t_0}(T_0(t,h),h)}, \tag{B-5}$$

$$\delta T(t_0,h) = -\frac{2h^2\delta u(t_0))}{T(t_0,h)}.$$

Combining the above equation with equation B-2 and A-5 gives

$$\delta p(t,h) = \frac{4hT_0(t,h)\delta u(T_0(t,h))}{t^2}\frac{\partial T_0}{\partial t}(t,h). \tag{B-6}$$

Combine equation A-6 with A-3 to obtain

$$\delta J[v] = \int dh \int dt\ \frac{\partial T_0}{\partial t}(t,h)\frac{4hT_0(t,h)\delta u(T_0(t,h))}{t^2}\frac{\partial d}{\partial t}(t,h)\left(\frac{\partial d}{\partial h} + p\frac{\partial d}{\partial t}\right). \tag{B-7}$$

Since

$$\delta u(t_0) = \frac{-2u^2(t_0)}{t_0}\int_0^{t_0} d\tau_0 v(\tau_0)\delta v(\tau_0),$$

then the gradient is

$$\nabla J[v](\tau) = -\int dh\ v(\tau)\int_{T(\tau,h)}^{t_{max}} dt\left\{\frac{\partial T_0}{\partial t}(t,h)\frac{p^2(t,h)}{2h}\frac{\partial d}{\partial t}(t,h)\left(p\frac{\partial d}{\partial t} + \frac{\partial d}{\partial h}\right)(t,h)\right\}. \tag{B-8}$$

Bibliography

[Barry and Kneale, 1980] Barry, K. Cavers, D. and C. Kneale, 1980, Seg-y - recommended standards for digital tape formats: Technical report, Society of Exploration Geophysics, Tulsa.

[Chauris and Noble, 2001] Chauris, H. and M. Noble, 2001, Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications to synthetic and real data sets: Geophysical Journal International, **144**, 14–26.

[de Hoop et al., 2005] de Hoop, M., S.-K. Foss, and B. Ursin, 2005, Depth-consistent reflection tomography using PP and PS seismic data: Geophysics, **70**, U51–U65.

[Dix, 1955] Dix, C., 1955, Seismic velocities from surface measurements: Geophysics, **20**, 68–86.

[Gauthier et al., 1986] Gauthier, O., A. Tarantola, and J. Virieux, 1986, Two-dimensional nonlinear inversion of seismic waveforms: Geophysics, **51**, 1387–1403.

[Kirkpatrick et al., 1983] Kirkpatrick, S., C. Gelatt, and M. Vecchi, 1983, Optimiza-

tion by simulated annealing: Science, **220**, 671–680.

[Kolb et al., 1986] Kolb, P., F. Collino, and P. Lailly, 1986, Prestack inversion of a 1D medium: Proceedings of IEEE 74, 498–506.

[Li and Symes, 2007] Li, J. and W. Symes, 2007, Interval velocity estimation via nmo-based differential semblance: Geophysics, **72**, U75–U88.

[Lyell, 1830] Lyell, C., 1830, Principles of Geology.

[Minkoff and Symes, 1997] Minkoff, S. and W. Symes, 1997, Full waveform inversion of marine reflection data in the plane-wave domain: Geophysics, **62**, 540–553.

[Mulder and ten Kroode, 2002] Mulder, W. and A. ten Kroode, 2002, Automatic velocity analysis by differential semblance optimization: Geophysics, **67**, 1184–1191.

[Neidell and Taner, 1971] Neidell, N. and T. Taner, 1971, Semblance and other coherency measures for multichannel data: Geophysics, **36**, 498–509.

[Nodedal and Wright, 1999] Nodedal, J. and S. Wright, 1999, Numerical optimization: Springer Verlag.

[Padula et al., 2004] Padula, A., S. Scott, and W. Symes, 2004, https://cvs.caam.rice.edu:3129/svn/symes/trunk/packs/rvlpack/.

[Sen and Stoffa, 1991a] Sen, M. and P. Stoffa, 1991a, Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane wave seismograms: Geophysics, **56**, 1794–1810.

[Sen and Stoffa, 1991b] ———, 1991b, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: Geophysics, **56**, 1624–1636.

[Stockwell and Cohen, 2004] Stockwell, J. and J. Cohen, 2004, Cwp/su: Seismic unix release no.37: a free package for seismic research and processing: Technical report, Center for Wave Phenomena, Colorado School of Mines.

[Symes, 1993] Symes, W., 1993, A differential semblance criterion for inversion of multioffset seismic reflection data: Journal of Geophysical Research, **98**, 2061–2073.

[Symes, 1998a] ———, 1998a, Mathematical foundations of reflection seismology, *in* Lecture notes.

[Symes, 1999] ———, 1999, All stationary points of differential semblance are asymptotic global minimizers: layered acoustics: Technical Report 99-26, CAAM, Rice University.

[Symes, 2008] ———, 2008, Migration velocity analysis and waveform inversion: Geophysical prospecting, **56**, 765–790.

[Symes and Carazzone, 1991] Symes, W. and J. Carazzone, 1991, velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

[Symes and Gockenbach, 1995] Symes, W. and M. Gockenbach, 1995, Waveform inversion for velocity: Where have all the minima gone?: 65th Annual International

Meeting of the Society of Exploration geophysicists, Expanded Abstracts, 1235–1239.

[Symes, 1998b] Symes, W. W., 1998b, High frequency asymptotics, differential semblance, and velocity analysis: 68th Annual International Meeting, Expanded Abstracts, 1616–1619, Society of Exploration Geophysicists.

[Toldi, 1989] Toldi, J., 1989, Velocity analysis without picking: Geophysics, **54**, 191–199.

[Winslow, 2000] Winslow, N., 2000, Joint inversion using the convolutional model: Master's thesis, CAAM, Rice University, Houston, Texas.

[Yilmaz, 1987] Yilmaz, O., 1987, Seismic data analysis: processing, inversion, and interpretation of seismic data: Society of exploration geophysics.